IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:

Francis J. Carr — Examiner: Teresa D. Wessendorf

Serial No.: 09/937,100 — Group Art Unit: 1639

Filed: September 20, 2001

Title: PROTEIN ISOLATION AND ANALYSIS

APPEAL BRIEF

Mail Stop: Appeal Brief Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

Further to the Notice of Appeal filed February 24, 2005, herewith is Appellant's Brief on Appeal. A check for the statutory fee of $500.00 fee for filing an Appeal Brief is enclosed. This is an appeal from the Final Rejection of August 25, 2004, a reply to which filed on January 25, 2005 and April 25, 2005 were not considered persuasive by the Supervisory Primary Examiner in an Office Communication of February 28, 2005.

(1) REAL PARTY IN INTEREST

The real party in interest is Biovation Limited as indicated in the assignment recorded September 20, 2001 (Real/Frame, 012293/0929).

(2) RELATED APPEALS AND INTERFERENCE

There are no related appeals or interferences known to appellant, the appellant's legal representative, or assignee which will directly affect or be directly affected by or have a bearing on the Board's decision in the pending appeal.

(3) STATUS OF THE CLAIMS

During the prosecution of the instant application, claims 1-74 were originally presented for examination. Claims 1-15 were cancelled, and Claims 16-51 are withdrawn. Claims 52-74 remain pending. A copy of all the pending claims is presented in the Appendix.

(4) STATUS OF AMENDMENTS AFTER FINAL

The amendment filed January 25, 2005 subsequent to the final rejection was entered and deemed to overcome the rejections under §112, second paragraph, §102, and §103. See, Advisory Action mailed January 25, 2005. The amendment April 25, 2005 subsequent to the final rejection was not entered and is not relied upon for the purposes of this Appeal.

(5) SUMMARY OF THE INVENTION

Claims 64-74 are directed to methods of identifying a protein from a library of individual proteins that binds to a target of interest. Claims 52-63 are directed to the libraries, per se. The claimed method recites certain structural claimed aspects, including "protease sensitive sites" and "identifier sequence tracts which are unique" possessed by the individual proteins in the library. The identifier sequence tracts can be used to recover ("identify") a protein in the library which has the desired binding characteristics after cleavage at the protease sensitive site. See, e.g., Claim 65(iv) and Claim 75(iv). Specification, Page 2, lines 25-30; Page 3, lines 23-29; Page 4, lines 16-40. The claimed method is generally useful for identifying proteins that possess an activity of interest. The term "barcode" is also used in the specification to refer to the "identifier

sequence tracts." Specification, Page 2 lines 21-29. The terms are used interchangeably in the arguments below.

(6) ISSUES

There are no prior art rejections. The sole issues remaining in this application are:

Whether the claims are unpatentable for lack of written description under §112, first paragraph and utility under §101. These grounds are treated together for the purposes of this Appeal since the issues are intertwined, and if the §112 issue is overcome, the §101 would fall, as well.

(7) GROUPING OF THE CLAIMS

Claims 52-56 and 58-74 stand or fall together. Claim 57 does not stand or fall apart from the others since it recites specific protease cleavages site for enterokinase and Factor Xa that have specific structural characteristics. See, specification, Page 3, lines 7-8.

(8) APPELLANTS' ARGUMENTS

It is alleged by the Patent Office that: "The claims or the specification does not recite for any specific structure of the library." See, Office action dated August 25, 2004, Page 4. This is not correct. Claim 52 and others clearly recite specific structural features, e.g., "individual identifier sequence amino acid tracts which are unique to said individual protein when bound to the specific target of interest, and are flanked by one or more protease sensitive sites ... " (Underlining added.) In Claims 65 and 74, e.g., these structures (i.e., identifiers and protease sensitive sites) are utilized to identify a protein with desired binding characteristics. This useful, novel, and unobvious method can be applied to many different kinds of proteins, including antibodies (e.g., Specification, Page 1, lines 12) and other binding moieties. A utility of the claimed invention involves the use of the identifiers and protease sensitive sites to select proteins with certain characteristics. The focus on the particular proteins to which the identifiers and protease sensitive sites are joined is misplaced. The same reasoning would preclude any general

method from being patented.

Identifier sequence amino acid tracts (also known as "barcodes") are fully described in the specification. The barcodes are amino acid tracts encoded for by oligonucleotides that are attached, in open reading frame, to the nucleotide sequence encoding the polypeptide of interest. A protease sensitive-site is placed between them to facilitate release of the peptide encoded by the barcode sequence. The polypeptide is expressed and then selected on the basis of an activity, e.g., a binding activity. Those polypeptides which possess the activity of interest can be then subjected to protease cleavage to cause the release of the barcode peptide. The peptide can be directly sequenced or analyzed by mass spectrometry to determine its amino acid sequence. This information is reversed translated into an oligonucleotide sequence that can be utilized to specifically amplify (by polymerase chain reaction) the polypeptide possessing the activity of interest. See, e.g., Specification, Page 2, line 25-Page 30, line 30.

The design of identifier sequence amino acid tracts (i.e., barcodes), including their structures, is described throughout the specification. For example, on Page 4, lines 23-44, an eight-amino acid barcode sequence using 17 of the 20 natural amino acids is described. A specific example is provided of a family of peptide barcodes and their corresponding oligonucleotide sequences. See, Page 5, lines 1-35. Example 2, beginning on Page 29 of the specification, provides a specific working example of barcode sequence. See, Page 30, lines 5-15. Another example is disclosed on Page 35, lines 30-35. Thus, the specification provides adequate written description of the barcode sequence, including specific structural examples.

Protease sensitive sites were well-known in the scientific community at the time the application was filed. The specification describes the structures for enterokinase, Factor Xa, and thrombin cleavage sites (which are also widely used in molecular biology). See, e.g., Page 3, lines 5-10 and Page 23, line 34-Page 24, line 2. A V8 protease cleavage site is also disclosed. See, Page 30, lines 5-15. Thus, specific structures of protease-sensitive sites are described in the specification. An Appellant is not required to disclose every species encompassed by their claims. *In re Angstadt*, 537 F.2d 498, 502-503, 190 USPQ 214, 218 (CCPA 1976). Other protease sites were known in the art. These include, e.g., enterokinase (Hopp et al., 1988), thrombin (Eaton et al., 1986; Manoharan et al., 1997), collagenase (Gehring et al., 1995), tobacco

etch virus (U.S. Pat. No. 5,532,142), IgA protease (U.S. Pat. No. 5,427,927), and dipeptide-specific proteases, such as aminopeptide B and carboxypeptides B, E, and N (U.S. Pat. No. 5,506,120). See, Exhibit 1 for a copy of the non-patent publications. A patent need not teach, and preferably omits, what is well known in the art. *In re Buchner*, 929 F.2d 660, 661, 18 USPQ2d 1331, 1332 (Fed. Cir. 1991); *Hybritech, Inc. v. Monoclonal Antibodies, Inc.*, 802 F.2d 1367, 1384, 231 USPQ 81, 94 (Fed. Cir. 1986), *cert. denied*, 480 U.S. 947 (1987); and *Lindemann Maschinenfabrik GMBH v. American Hoist & Derrick Co.*, 730 F.2d 1452, 1463, 221 USPQ 481, 489 (Fed. Cir. 1984).

It is not correct that all functional descriptions of genetic material do not meet the written description requirements. This was expressly stated in *Enzo Biochem Inc.* v. *Gen-Probe*, 63 USPQ2d 1609, 1613 (Fed. Cir. 2002). As pointed out by the *Enzo* Court, antibody claims are considered to be in compliance with §112, first paragraph, even though no structural features of the antibody are disclosed in the specification. Similarly, the identifier and protease sites are sufficient to comply with the written description requirements. (Indeed, the present examiner has allowed claims reciting "exonuclease sensitive site" which is analogous to the feature recited in the pending claims that are now alleged to lack written description. See, U.S. Pat. No. 6,322,969).

The examiner's reasoning in relying on *Eli Lilly* and *Fiers* on Page 7 of the Office action dated August 25, 2004, is neither relevant nor appropriate. In those cases, the Appellant had discovered a single sequence for a protein from one mammalian species, but was attempting to cover the entire genus of mammalian proteins. This is not the case here. Appellant is claiming a general method of identifying proteins, and a library that is useful in this claimed method, not a specific protein as in the *Eli Lilly* and *Fiers* cases. Appellant is not claiming beyond what is described and enabled in the specification, i.e., the general method and use of a library having certain structural features that make it useful in the method.

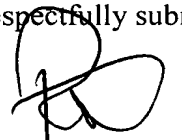
Claims of such type and scope have been granted by the Patent Office (e.g., System to detect protein-protein interactions, U.S. Pat Nos. 5,283,173; 5,468,617; and 5,667,973). There is no statutory reason to preclude it.

The allegation on Page 5 of the Office action that " ... there was no immediately apparent [sic] or 'real world' utility as of the filing date" is preposterous. Not only does the specification provide a number of uses of the claimed library and methods, but beginning on Page 26 of the Specification, several actual and working examples are described in which proteins with a binding affinity to a target were identified. This clearly is a substantial, specific, and credible utility.

(9) Conclusion

In view of the arguments and authorities presented above, Appellants request that the Examiner's action in making and maintaining the rejection under 35 USC §101 and §112, first paragraph, be reversed and that the application be allowed.

Respectfully submitted,



Richard M. Lebovitz (Reg. No. 37,067)
Attorney for Appellant(s)
MILLEN, WHITE, ZELANO & BRANIGAN, P.C.
Arlington Courthouse Plaza I, Suite 1400
2200 Clarendon Boulevard
Arlington, Virginia 22201
(703) 812-5317 [Direct Dial]
(703) 243-6410 [Facsimile]
Internet Address: lebovitz@mwzb.com

Filed: June 24, 2005

APPENDIX

Claim 1 **(Cancelled)**

Claim 2 **(Cancelled)**

Claim 3 **(Cancelled)**

Claim 4 **(Cancelled)**

Claim 5 **(Cancelled)**

Claim 6 **(Cancelled)**

Claim 7 **(Cancelled)**

Claim 8 **(Cancelled)**

Claim 9 **(Cancelled)**

Claim 10 **(Cancelled)**

Claim 11 **(Cancelled)**

Claim 12 **(Cancelled)**

Claim 13 **(Cancelled)**

Claim 14 **(Cancelled)**

Claim 15 **(Cancelled)**

Claim 16 **(Withdrawn)** A method of screening a protein library comprising screening said library for one or more desired properties, followed by dereplication to identify one or more individual proteins in the library having the desired property.

Claim 17 **(Withdrawn)** A method as claimed in claim 16 wherein the library is screened for binding to a target moiety.

Claim 18 **(Withdrawn)** A method as claimed in claim 17 wherein binding is detected by mass spectrometry, particularly matrix-assisted laser desorption/ionization time-of-flight (MALDI-ToF) spectrometry.

Claim 19 **(Withdrawn)** A method as claimed in claim 16 wherein the library is screened for a specific biological activity.

Claim 20 **(Withdrawn)** A method as claimed in claim 17 wherein the target is a complex mixture, eg a mixture of molecules, whole cells or cell membranes.

Claim 21 **(Withdrawn)** A method of protein identification and/or sequencing comprising providing a library of individual proteins, one or more of which may bind to a target of interest, wherein each individual protein, together with its gene, is bound to an "associating moiety".

Claim 22 **(Withdrawn)** A method as claimed in claim 21 wherein the library of proteins is brought into contact with the target of interest either before or after the "associating moiety".

Claim 23 **(Withdrawn)** A method as claimed in claim 21 wherein after screening for binding to the target the library is dereplicated to identify one or more proteins with a desirable property, proteins which bind to the target.

Claim 24 **(Withdrawn)** A method as claimed in claim 21 where the "associating moiety" is a particle.

Claim 25 **(Withdrawn)** A method as claimed in claim 24 wherein the particle is a latex bead.

Claim 26 **(Withdrawm)** A method as claimed in claim 21 wherein the "associating moiety" is a protein or protein complex.

Claim 27 **(Withdrawn)** A method as claimed in claim 26 wherein the "associating moiety" is avidin or steptavidin and each of the proteins in the library and their associated genes are biotinylated.

Claim 28 **(Withdrawn)** A method as claimed in claim 21 wherein the associating moiety is a bispecific binding molecule capable of binding to both the proteins and genes.

Claim 29 **(Withdrawn)** A method as claimed in claim 21 wherein the associating moiety is a living cell or cellular virus such as a bacteria or bacteriophage.

Claim 30 **(Withdrawn)** A method as claimed in claim 21 wherein one or other molecules which alter the properties of the proteins in the library are bound to the associating moiety

Claim 31 **(Withdrawn)** A method as claimed in claim 21 wherein the genes encoding the proteins in the library are are attached to the associating moiety rior to synthesis of the individual proteins.

Claim 32 **(Withdrawn)** A method as claimed in claim 21 wherein the library of proteins is a library of antibody proteins, eg a library of antibody domains such as Fvs.

Claim 33 **(Withdrawn)** A method of protein identification and/or sequencing comprising providing a library of individual proteins, one or more of which may bind to a target of interest, wherein each individual protein is attached to an individual "coding moiety".

Claim 34 **(Withdrawn)** A method as claimed in claim 33 wherein the "coding moieties" are particles with unique identifier "codes".

Claim 35 **(Withdrawn)** A method as claimed in claim 34 wherein the "codes" are different ratios of measurable signal, eg fluorescent, chemiluminescent or radioactive labels, or a physical feature such as a unique marking.

Claim 36 **(Withdrawn)** A method for analyzing mixtures of proteins comprising:

(iii) digestion or cleavage of the protein mixture;

(iv) fractionation of the resultant peptides; and

(v) analysis of the resultant peptides by means of their mass and/or sequence.

Claim 37 **(Withdrawn)** A method as claimed in claim 36 wherein the fractionation in step (ii) is carried out using a library of protein binding agents.

Claim 38 **(Withdrawn)** A method as claimed in claim 36 wherein the resultant peptides are subjected to physical fractionation and/or chemical tagging as part of the fractionation step.

Claim 39 **(Withdrawn)** A method as claimed in claim 36 wherein the resultant peptides are subjected to addition of one or more amino acids as part of the fractionation step.

Claim 40 **(Withdrawn)** A method as claimed in claim 37 wherein the library of protein binding agents is a library of antibodies or antibody fragments.

Claim 41 **(Withdrawn)** A method as claimed in claim 37 wherein the protein binding agents are major histocompatibility proteins, T cell receptors and natural proteins or protein domains involved in protein-protein binding interactions, such as SH1 domains.

Claim 42 **(Withdrawn)** A method as claimed in claim 40 wherein the library of protein binding agents is pre-selected for binding to one or more proteins or peptides derived from the protein mixture or a related protein mixture under analysis.

Claim 43 **(Withdrawn)** A method as claimed in claim 42 wherein the protein mixture is derived from a normalised recombinant gene library.

Claim 44 **(Withdrawn)** A method as claimed in claim 36 wherein the protein mixture is initially bound to a solid phase prior to digestion or cleavage either via the N or C-terminus or via specific amino acids or via specific sequences of amino acids.

Claim 45 **(Withdrawn)** A method as claimed in claim 36 wherein specific amino acids or modified amino acids found in the proteins are derivatised prior to binding to a solid phase, such binding occurring either before or after digestion or cleavage of the protein mixtures.

Claim 46 **(Withdrawn)** A method as claimed in claim 45 wherein the specific, or modified amino acids are derivatised with biotin prior to binding to avidin or streptavidin.

Claim 47 **(Withdrawn)** A method as claimed in claim 45 wherein specific, or modified, amino acids are derivatised with ligands prior to binding to ligand-specific affinity reagents.

Claim 48 **(Withdrawn)** A method as claimed in claim 36 wherein specific naturally modified amino acids found in the proteins are bound to a solid phase using modification specific affinity reagents, such binding occurring either before or after digestion or cleavage of the protein mixtures.

Claim 49 **(Withdrawn)** A method as claimed in claim 45 wherein more than one cycle of digestion/cleavage and derivatisation is carried out.

Claim 50 **(Withdrawn)** A method as claimed in claim 49 wherein mass analysis is carried out after each cycle of digestion or cleavage.

Claim 51 **(Withdrawn)** A method as claimed in claim 36 wherein peptides released after digestion/cleavage are fractionated using physical methods such as HPLC before or after fractionation using protein binding agents.

Claim 52 **(Previously Presented)** A library of individual proteins, one or more of which being able to bind to a target of interest, each of said proteins comprising within its amino acid sequence, or terminal to it, one or more individual identifier sequence amino acid tracts which are unique to said individual protein when bound to the specific target of interest, and are flanked by one or more protease sensitive sites, said library thus providing proteins comprising a diversity of said individual identifier sequence tracts.

Claim 53 **(Previously Presented)** A library of claim 52, wherein said identifier sequence tracts have been generated randomly or semi-randomly.

Claim 54 **(Previously Presented)** A library of claim 52, wherein an individual protein of said library comprises multiple identifier sequence tracts.

Claim 55 **(Previously Presented)** A library according claim 54, wherein an individual protein comprises two adjacent identifier sequence tracts.

Claim 56 **(Previously Presented)** A library according to claim 52, wherein said protease sensitive site is a recognition site for endoprotease digestion.

Claim 57 **(Previously Presented)** A library of claim 56, wherein the recognition site is the site for enterokinase or Factor Xa.

Claim 58 **(Previously Presented)** A library according to claim 52, wherein said target of interest is a protein.

Claim 59 **(Previously Presented)** A library of claim 52, wherein the individual proteins of the library are antibodies or recombinant antibody domains, which comprise antibody variable regions.

Claim 60 **(Previously Presented)** A library of claim 59, wherein the antibody domain is an Fv domain consisting of a VH and a VL chain.

Claim 61 **(Previously Presented)** A library according to claim 60, wherein each chain has its own identifier sequence tract.

Claim 62 **(Previously Presented)** A library according to claim 59, wherein said identifier sequence tract is C-terminal to the Fv domain sequence.

Claim 63 **(Previously Presented)** A library according to 59, wherein the target of interest is an antigen.

Claim 64 **(Previously Presented)** A method of identifying a protein which binds to a target of interest, wherein the protein is a member of the library of claim 52, comprising:
(i) bringing each of the individual proteins of the library comprising said one or more individual identifier sequence tracts and said protease sensitive site(s) in contact with one or more of said targets of interest under conditions for an individual protein of said library to bind to said target of interest to form a complex,
(ii) isolating the complex formed by the individual protein and the target of interest,

Claim 65 **(Previously Presented)** A method of claim 64, further comprising:
(iii) digesting said complex to cleave said protease sensitive sites releasing said individual identifier sequence tract(s), and
(iv) determining said released sequence tract(s) by which said individual protein is finally identifiable.

Claim 66 **(Previously Presented)** A method of claim 64, wherein the individual protein has been encoded by a DNA construct comprising nucleic sequences coding for said one or more individual identifier sequence tracts and said protease sensitive site(s).

Claim 67 **(Previously Presented)** A method of claim 66, wherein said DNA construct is part of a vector system.

Claim 68 **(Previously Presented)** A method according to claim 64, wherein the individual protein and target of interest bind in solution.

Claim 69 **(Previously Presented)** A method according to claim 64, wherein the complex of protein / target is isolated by removal of non-bound molecules.

Claim 70 **(Previously Presented)** A method according to claim 65, wherein digestion of said complex is achieved by endoprotease.

Claim 71 **(Previously Presented)** A method according to claim 65, wherein determination of the released identifier sequence tract(s) is achieved by mass spectrometry.

Claim 72 **(Previously Presented)** A method of claim 71, wherein the mass spectrometry is MALDI-ToF.

Claim 73 **(Previously Presented)** A method according to claim 64, wherein the individual proteins of the library are antibodies or antibody domains having a variable region and the targets of interest are antigens.

Claim 74 **(Previously Presented)** A method of identifying a protein which binds to a target of interest wherein the protein is a member of the library as defined in claim 52, comprising the following steps:

(i) bringing each of the individual proteins of the library comprising said one or more individual identifier sequence tracts and said protease sensitive site(s) in contact or association with one or more of said targets of interest,

(ii) isolating the complex formed by the individual protein and the target of interest,

(iii) digesting said complex to cleave the introduced protease sensitive sites releasing said individual identifier sequence tract(s), and

(iv) determining said released sequence tract(s) and using such sequence information to recover the individual protein library member from the library, and

(v) including one or more further rounds of screening and enrichment for a protein which binds to the target of interest.

EXHIBIT

1

Mao, D., & Wallace, B. A. (1984) *Biochemistry 23*, 2667–2673.
Nathans, J., & Hogness, D. S. (1983) *Cell (Cambridge, Mass) 34*, 807–814.
Oesterhelt, D., & Stoeckenius, W. (1974) *Methods Enzymol. 31*, 667–678.
Ogurusu, T., Maeda, A., Sasaki, N., & Yoshizawa, T. (1982) *Biochim. Biophys. Acta 682*, 446–451.
Rehorek, M., & Heyn, M. P. (1979) *Biochemistry 18*, 4977–4983.
Schobert, B., & Lanyi, J. K. (1982) *J. Biol. Chem. 257*, 10306–10313.
Schobert, B., Lanyi, J. K., & Cragoe, E. J., Jr. (1983) *J. Biol. Chem. 258*, 15158–15164.
Smith, S. O., Marvin, M. J., Bogomolni, R. A., & Mathies, R. A. (1984) *J. Biol. Chem. 259*, 12326–12329.
Steiner, M., Oesterhelt, D., Ariki, M., & Lanyi, J. K. (1984) *J. Biol. Chem. 259*, 2179–2184.
Stoeckenius, W., & Bogomolni, R. A. (1982) *Annu. Rev. Biochem. 51*, 587–616.
Sugiyama, Y., & Mukohata, Y. (1984) *J. Biochem. (Tokyo) 96*, 413–420.
Taylor, M. E., Bogomolni, R. A., & Weber, H. J. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 6172–6176.

Proteolytic Processing of Human Factor VIII. Correlation of Specific Cleavages by Thrombin, Factor Xa, and Activated Protein C with Activation and Inactivation of Factor VIII Coagulant Activity

Dan [redacted] Henry Rodriguez, and Gordon A. Vehar*
Department of Protein Biochemistry, Genentech, Inc., South San Francisco, California 94080
Received June 11, 1985

ABSTRACT: Human factor VIII was isolated from commercial factor VIII concentrates and found to consist of multiple polypeptides with molecular weights ranging from 80 000 to 210 000. Immunological and amino acid sequence data identified these polypeptides as subunits of factor VIII. N-Terminal amino acid sequence analysis determined that the M_r 210 000 and 80 000 proteins are derived from the N- and C-terminal portions of factor VIII, respectively; M_r 90 000–180 000 polypeptides are derived from the M_r 210 000 polypeptide by C-terminal cleavages. Treatment of purified factor VIII with thrombin resulted in proteolysis of M_r 80 000–210 000 proteins and the generation of polypeptides of M_r 73 000, 50 000, and 43 000. Maximum coagulant activity of thrombin-activated factor VIII was correlated with the generation of these polypeptides. The proteolysis as well as activation of factor VIII by thrombin was found to be markedly dependent on $CaCl_2$ concentration. Proteolysis of factor VIII with activated protein C (APC) resulted in degradation of the M_r 90 000–210 000 proteins with the generation of an M_r 45 000 fragment. This cleavage correlated with inactivation of factor VIII by APC. The M_r 80 000 protein was not degraded by APC. Factor Xa cleaved the M_r 80 000–210 000 factor VIII proteins, resulting in the generation of fragments of M_r 73 000, 67 000, 50 000, 45 000, and 43 000. Factor Xa was found to initially activate and subsequently inactivate factor VIII. Activation by factor Xa correlated with the generation of M_r 73 000, 50 000, and 43 000 polypeptides while inactivation correlated with the cleavage of M_r 73 000 and 50 000 polypeptides to fragments of M_r 67 000 and 45 000, respectively. The cleavage sites in factor VIII of thrombin, factor Xa, and APC were identified by amino acid sequencing of the fragments generated after cleavage of factor VIII by these proteases. Interestingly, factor Xa was found to cleave factor VIII at the same sites as APC and thrombin. This may explain why factor Xa activates as well as inactivates factor VIII.

Purification of factor VIII (antihemophilic factor) from plasma indicates that its coagulant activity is associated with multiple polypeptide chains having molecular weights ranging from 80 000 to 210 000 (Vehar & Davie, 1980; Fass et al., 1982; Fulcher & Zimmerman, 1982; Rotblat et al., 1985). Recently, cDNA clones encoding the entire factor VIII protein sequence have been obtained (Toole et al., 1984; Wood et al., 1984). The amino acid sequence deduced from such clones predicts a mature single-chain protein (2332 amino acids) having a molecular weight of ~300 000 (Wood et al., 1984; Toole et al., 1984). Sequence data obtained from the protein chains of purified factor VIII preparations have been shown to overlap with the sequence predicted from the cDNA clones (Toole et al., 1984; Vehar et al., 1984), and the purification of a single-chain precursor having a M_r >300 000 has been reported (Rotblat et al., 1985). Thus, if factor VIII circulates in plasma as a single-chain form, it is partially degraded during its purification, yielding a form with multiple polypeptide chains.

Amino acid sequence analyses also revealed the orientation of the protein chains associated with factor VIII to the single-chain precursor deduced from the cDNA sequence (Vehar et al., 1984; Toole et al., 1984). Such data show that the M_r 210 000 and 80 000 proteins represent the N-terminal and C-terminal portions of factor VIII, respectively (Vehar et al., 1984; Toole et al., 1984). It is proposed that several proteolytic cleavages on the C-terminal side of the M_r 210 000 protein generate a series of proteins with molecular weights between 90 000 and 180 000 (Vehar et al., 1984; Toole et al., 1984).

Recently, thrombin activation of factor VIII coagulant activity has been shown to be associated with specific proteolysis of factor VIII protein chains (Vehar & Davie, 1980; Fass et al., 1982; Fulcher et al., 1983, 1984; Loller et al., 1984; Rotblat et al., 1985). During thrombin activation of purified

0006-2960/86/0425-0505$01.50/0

human factor VIII, proteins with M_r 110000–210000 appear to be proteolyzed to generate an M_r 90000 protein, while the M_r 80000 protein is cleaved to an M_r 73000 fragment (Fulcher et al., 1983). Fulcher et al. (1983) suggest that the M_r 90000 and 73000 proteins are the active subunits of thrombin-activated factor VIII and that cleavage of the M_r 90000 protein by thrombin (yielding fragments of M_r ~50000 and ~43000) inactivates factor VIII coagulant activity. In contrast, studies with purified porcine factor VIII suggest that cleavage of an M_r 82000 protein, which is analogous to the M_r 90000 moiety of human factor VIII, results in further activation of factor VIII coagulant activity (Fass et al., 1982; Loller et al., 1984). Other than species differences, the reason(s) for this discrepancy is (are) unknown.

The inactivation of human factor VIII by activated protein C (APC), a vitamin K dependent plasma protease, has also been correlated with limited proteolysis of the factor VIII protein (Fulcher et al., 1984). Cleavage of factor VIII with APC results in the proteolysis of the M_r 90000–210000 proteins with the concomitant appearance of an M_r 45000 fragment (Fulcher et al., 1984).

The above-mentioned studies clearly show that specific proteolytic processing of factor VIII regulates factor VIII coagulant activity. In this report, we compare the effects of thrombin, factor Xa [also known to activate factor VIII (Vehar & Davie, 1980; Davie et al., 1975)], and APC on factor VIII coagulant activity and correlate the changes in activity with changes in factor VIII subunit structure. Furthermore, most of the cleavage sites of these proteases have been identified by amino acid sequence analyses of the fragments generated by the proteolysis of factor VIII. Knowledge of these sites not only allows the cleavage patterns of these proteases to be compared but also begins to illustrate the basis of the mechanisms that alter factor VIII coagulant activity.

Materials and Methods

Human factor Xa, human activated protein C (APC), and human α-thrombin were all generous gifts from Dr. Walter Kisiel (The University of New Mexico). Affi-gel 10 was from Bio-Rad; rabbit brain cephalin and phenylmethanesulfonyl fluoride (PMSF) were from Sigma Chemical Co.; Platelin was obtained from General Diagnostics; factor VIII deficient and normal human plasmas were from George King Biomedical; factor VIII chromogenic Coatest assay was from Helena. Bio-Gel A-15m void volume fractions enriched in factor VIII/von Willebrand factor (vWF) complexes were prepared from commercial concentrates and were a generous gift of Cutter Laboratories and Dr. D. Schroeder.

Purification of Human Factor VIII. Commercial factor VIII concentrate from Cutter Laboratories was resolved on a Bio-Gel A-15m column as described by Fay et al. (1982). The V_0 fraction containing factor VIII coagulant activity was made 1 mM PMSF and 35 mM β-mercaptoethanol. This results in the reduction of von Willebrand factor (vWF)/factor VIII complexes, which has been shown to cause their dissociation without significantly affecting factor VIII coagulant activity (Vehar & Davie, 1980; Savidge et al., 1979). Also, the functional and structural properties of the factor VIII preparations isolated here are very similar to factor VIII preparations isolated by others in the absence of reducing agents (Fulcher & Zimmerman, 1982; Rotblat et al., 1985). The reduced V_0 was batch-separated with DEAE-Sepharose that had been equilibrated in a 0.02 M imidazole, pH 6.9, buffer containing 0.15 M NaCl, 0.01 M $CaCl_2$, 0.02 M glycine hydrochloride ethyl ester, 5% glycerol, and 1 mM PMSF (VIII buffer). Twenty milliliters of DEAE-Sepharose was added for every liter of V_0 fraction. After being stirred for 2–3 h at 4 °C, the resin was poured into a column and washed with 5 column volumes of VIII buffer. Factor VIII was step-eluted with VIII buffer containing 0.11 M $CaCl_2$. A factor VIII monoclonal antibody column was prepared by coupling 10 mg of factor VIII monoclonal antibody to 2 mL of Affi-gel 10 (Wood et al., 1984). The resulting column was equilibrated in 0.05 M imidazole, pH 6.9, buffer containing 0.15 M NaCl, 0.01 M $CaCl_2$, 5% glycerol, and 1 mM PMSF. The factor VIII DEAE pool was applied to the antibody column, and the column was washed with 50 column volumes of the above buffer. Factor VIII was eluted with the same buffer containing 1.0 M KI. Fractions containing factor VIII activity were pooled and dialyzed against 0.05 M tris(hydroxymethyl)-aminomethane (Tris), pH 7.5, 0.15 M NaCl, 2.5 mM $CaCl_2$, 5% glycerol, and 1 mM PMSF and stored at −70 °C. Factor VIII activity was measured either by coagulation analysis or by the factor VIII chromogenic Coatest assay as described by Wood et al. (1984). Protein concentration was determined by the method of Bradford (1976).

Cleavage of Factor VIII by Thrombin, Factor Xa, and APC. For N-terminal amino acid sequence analysis, approximately 0.5–1.0 mg of factor VIII was incubated with either thrombin, factor Xa, or APC at a 1/50 ratio (w/w). In the case of factor Xa and APC, $^1/_{10}$th sample volume of rabbit brain cephalin was included in the reaction as a source of phospholipid. After 1–2 h at 37 °C, the reaction was stopped by adding sodium dodecyl sulfate (SDS) to 0.4% and immediately heating the samples to 80 °C. Proteolyzed factor VIII was subsequently resolved on 5–10% polyacrylamide gradient gels in the presence of SDS [SDS–polyacrylamide gel electrophoresis (PAGE)]. Electrophoresis was carried out according to the method of Laemmli (1970). After staining with Coomassie blue, factor VIII peptides were excised and electroeluted according to the method of Hunkapiller et al. (1983). Gel-eluted peptides were subjected to N-terminal amino acid sequence analysis using an Applied Biosystems gas phase sequenator (Hewick et al., 1982) modified for on-line phenylthiohydantoin identification (H. Rodriguez, unpublished results).

For subunit and activity analysis during proteolysis, aliquots of factor VIII (110 μg/mL, 400–700 units/mL) in 0.05 M Tris, pH 7.5, 0.15 M NaCl, 2.5 mM $CaCl_2$, and 5% glycerol were incubated with either thrombin (1.5 μg/mL), factor Xa (2 μg/mL) or APC (4 μg/mL) for 0–120 min (thrombin, Xa) or 0–30 min (APC). Rabbit brain cephalin ($^1/_{10}$th sample volume) was added in reactions containing APC or factor Xa. At the end of each time point, a 10-μL aliquot of the reaction was removed, diluted into 0.05 M Tris, pH 7.3, containing 0.2% bovine serum albumin (BSA), and assayed by coagulation analysis. To the remainder of the aliquot was added SDS to 0.5%, and the sample was immediately heated to 80 °C. Proteolyzed factor VIII was subsequently resolved on 6–12% SDS–polyacrylamide gels. Proteins were visualized by silver staining (Morrissey, 1981).

Results

Purification of Factor VIII. Factor VIII was purified by initially resolving plasma concentrates on a Bio-Gel A-15m column as previously described (Fay et al., 1982). This allowed the partial purification of von Willebrand factor (vWF)/factor VIII complexes which elute in the void volume. These complexes were dissociated by reduction with β-mercaptoethanol and resolved by chromatography on DEAE-Sepharose. Factor VIII obtained from the DEAE-Sepharose chromatograph was subsequently purified to homogeneity by using a factor VIII monoclonal antibody column. A typical purification starting

Table I: Purification of Factor VIII

step	total act. (units)[a]	total protein[b] (mg)	sp act. (units/mg)	% recovery
5 L of V_0 (A-15m)	30000	2800	10	100
DEAE-Sepharose	15000	27.0	550	50
factor VIII monoclonal antibody column	8500	1.86	4500	28

[a] Activity was determined by coagulation analysis using human factor VIII deficient plasma. [b] Protein was measured by the method of Bradford (1976).

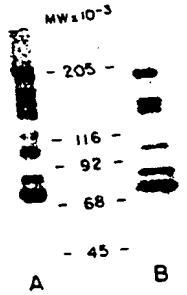


FIGURE 1: SDS–PAGE and western blot analysis of purified human factor VIII. Factor VIII (3–4 μg) was resolved on a 5–10% SDS–polyacrylamide gel and either silver-stained (A) or transferred to nitrocellulose for western blot analysis (B). Two different factor VIII monoclonal antibodies were used for western analyses. One cross-reacts with the M_r 90000–210000 polypeptides, while the other cross-reacts with the M_r 80000 polypeptide (Vehar et al., 1984; Wood et al., 1984).

with material obtained from the Bio-Gel A-15m column is shown in Table I. In this case, factor VIII was purified approximately 5000-fold over plasma concentrates and had a specific activity of 4500 units/mg.

When analyzed by SDS–PAGE, purified factor VIII was resolved into multiple protein chains having molecular weights ranging from 80000 to 210000 (Figure 1). This pattern of proteins is similar to that observed by others who have analyzed purified human factor VIII by SDS–PAGE (Fulcher & Zimmerman, 1982; Rotblat et al., 1985). When resolved under nonreducing conditions, this pattern remained unchanged (data not shown). Western blot analysis demonstrated that all the proteins associated with purified factor VIII cross-reacted with specific factor VIII monoclonal antibodies (Figure 1). Furthermore, as shown below, amino acid sequence analyses of these proteins, and comparison of these sequences with the factor VIII DNA sequence, demonstrate identity.

Presumably, each of the proteins of M_r 90000–210000 forms a complex (perhaps calcium linked) with the M_r 80000 subunit. This is evidenced by the purification of factor VIII consisting of M_r 80000–210000 proteins using a monoclonal antibody that recognizes only the M_r 80000 moiety. We have also found that after factor VIII is bound to a monoclonal antibody column specific for the M_r 80000 moiety, the M_r 90000–210000 proteins can be eluted with ethylendiaminetetraacetic acid (EDTA) (unpublished results). Similar results have been obtained for porcine factor VIII (Fass et al., 1982).

Proteolysis of Factor VIII by Thrombin, APC, and Factor Xa. N-Terminal amino acid sequence analysis of factor VIII proteins reveals that the M_r 90000–210000 proteins have the same N-terminal sequence, while that of the M_r 80000 protein (which sometimes appeared as a doublet) is distinct (Figure 3). Alignment of these sequences with the amino acid sequence deduced from the cDNA for factor VIII shows that the M_r 210000 and 80000 proteins represent the N-terminal and

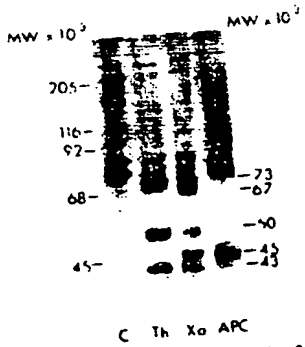


FIGURE 2: Cleavage of factor VIII by thrombin, factor Xa, and APC. Factor VIII (110 μg/mL, ~700 units/mL) was incubated for 1 h at 37 °C with either thrombin (1.5 μg/mL), factor Xa (1 μg/mL), or APC (2 μg/mL). In the case for factor Xa and APC, 1/10th volume of rabbit brain cephalin was included in the reaction. The reaction was stopped by the addition of SDS to 0.5% and heating to 80 °C. Proteins were subsequently resolved on a 6–12% SDS–polyacrylamide gel.

C-terminal portions of the factor VIII single-chain precursor, respectively (Figure 3; Toole et al., 1984; Vehar et al., 1984; Wood et al., 1984). Five polypeptide chains were routinely observed with M_r 110000–180000 (Figure 1). Presumably, C-terminal cleavage of the M_r 210000 protein generates these fragments. The sites that are cleaved to generate these five proteins are unknown as is the protease that makes them. It has been shown, however, that cleavage of the M_r 110000–210000 proteins by thrombin at position 740 generates the M_r 90000 protein (Toole et al., 1984).

Proteolysis of factor VIII by thrombin results in the degradation of the M_r 80000–210000 proteins and the appearance of polypeptides of M_r 73000, 50000, and 43000 (Figure 2). N-Terminal sequence analysis of the M_r 50000 and 43000 polypeptides shows that they are derived from the M_r 90000 protein by cleavage by position 372 (Figure 3; Vehar et al., 1984). The M_r 50000 and 43000 polypeptides represent the N-terminal and C-terminal portions of the M_r 90000 protein, respectively. The N-terminal sequence of the M_r 73000 polypeptide shows that it arises from the cleavage of the M_r 80000 protein at arginine-1689 (Figure 3; Vehar et al., 1984). This results in the possible removal of the N-terminal 44 amino acids of the M_r 80000 protein. This 44 amino acid polypeptide is acidic as it contains 15 Asp and Glu residues and only 4 Lys and Arg residues (Vehar et al., 1984). Like the M_r 80000 protein, the M_r 73000 polypeptide also appeared on SDS–PAGE as a doublet.

Factor Xa appears to proteolyze factor VIII more extensively than thrombin (Figure 2). Like thrombin, the M_r 80000–210000 proteins of factor VIII are all cleaved. However, in addition to the polypeptides of M_r 73000, 50000, and 43000, polypeptides of M_r 45000 and 67000 also appear after treatment of factor VIII with factor Xa (Figure 2). N-Terminal sequence analysis shows that the M_r 50000 and 43000 polypeptides originate from cleavage of the M_r 90000 protein at position 372 (Figure 3) as was observed with thrombin. Presumably, the M_r 90000 protein arises from cleavage of the M_r 110000–210000 proteins at position 740 by factor Xa. The M_r 45000 polypeptide has the same N-terminal sequence as the M_r 50000 and 90000–210000 polypeptides (Figure 3). The site at which this cleavage occurs has yet to be determined. However, on the basis of the size difference between the M_r 50000 and 45000 polypeptides, and since factor Xa is specific for arginine residues, cleavage at arginine-336 of the M_r 50000 or M_r 90000–210000 poly-

Polypeptide	No Protease	Thrombin	Factor Xa	APC
110,000-210,000	[A]ITRRYYLGAVELSWD			
90,000	ATRRYYL			
80,000	EITRTTLQSDQ			
73,000		ISFQKKITRHYFIAAVDERL	SFQKKTRHYF	
67,000			AQSGSVPQFKKVVFQ	
50,000		ATRRYYLGAVEL-W-YM	ATRRYYLGAVELS	
45,000			ATRRYYLGAV	ATRRYLGAVEL
43,000		ISIVAKK--E-WV-T	SVAKKHPK	

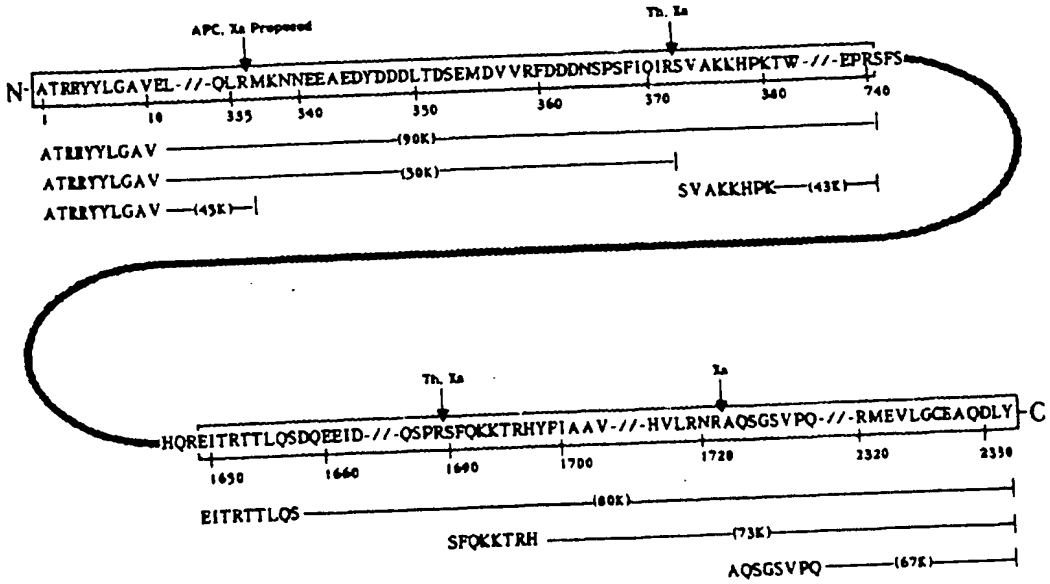


FIGURE 3: N-Terminal sequence of factor VIII polypeptides and their position within the factor VIII molecule. Factor VIII (0.5–1.0 mg) was proteolyzed with either thrombin, factor Xa, or APC, or not at all, and resolved on a 5–10% SDS–polyacrylamide gel. Subsequently, the polypeptides as shown were excised, gel-eluted, and subjected to N-terminal amino acid sequencing as described under Materials and Methods. The N-terminal amino acid sequence shown for the M_r 110 000–210 000 polypeptides was determined by pooling these peptides after gel elution. Only one sequence was obtained from this pool, and quantitation indicates that N-terminal blockage of these polypeptides was minimal (data not shown). Dashes indicate positions where no residue could be identified. Brackets indicate the tentative assignment for that amino acid. The factor VIII sequence shown in the boxes is deduced from the DNA sequence of factor VIII (Wood et al., 1984; Toole et al., 1984; Gitschier et al., 1984). Only that sequence which is necessary to show the orientation of the above-sequenced polypeptides is shown.

peptides may generate the M_r 45 000 fragment (Figure 3). The M_r 73 000 and 67 000 polypeptides are derived by cleavage of the M_r 80 000 protein at positions 1689 and 1721, respectively (Figure 3). The M_r 67 000 polypeptide also appears as a doublet on SDS–PAGE.

Cleavage of factor VIII by APC results in proteolysis of M_r 90 000–210 000 proteins with the appearance of an M_r 45 000 fragment (Figure 2). The M_r 80 000 protein is not cleaved by APC. These results are similar to those of Fulcher et al. (1984). The N-terminal sequence of the M_r 45 000 fragment was found to be the same as the N-terminal sequence of the M_r 90 000–210 000 proteins (Figure 3). Therefore, this fragment is derived from the N-terminal of factor VIII. The site at which APC cleaves factor VIII to generate this fragment has not been determined. This cleavage, however, appears to be the same cleavage made by factor Xa that also generates an M_r 45 000 fragment from the N-terminal of factor VIII. Cleavage at this site by factor Xa is not due to contaminating APC since antibodies against APC did not inhibit factor Xa from making this cleavage (data not shown). By SDS–PAGE, we could not reproducibly detect the C-terminal portions of the M_r 90 000–210 000 proteins after proteolysis of factor VIII by APC. In Figure 6, polypeptides with molecular weights of ~47 000, ~49 000, and 67 000 are apparent after APC cleavage. The appearance of these polypeptides, however, was only transitory. Also, only one sequence was observed when the M_r 45 000 fragment was sequenced.

Activation of Factor VIII by Thrombin. A time course treatment of factor VIII with catalytic amounts of thrombin resulted in a 36-fold increase in factor VIII coagulant activity (Figure 4). After maximum activity was reached, thrombin-activated factor VIII appeared to remain stable for at least 1 h at 37 °C. Figure 4 does show a slight decrease in activity at the 2-h time point; however, in other experiments, this decrease was not seen. Factor VIII that was not activated with thrombin remained stable throughout the 2-h time course (data not shown). Analysis of factor VIII subunit structure during thrombin activation shows that factor VIII coagulant activity dramatically increases with the generation of fragments of M_r 73 000, 50 000, and 43 000 (Figure 5). Thrombin, therefore, appears to activate factor VIII by initially cleaving M_r 110 000–210 000 proteins to generate the M_r 90 000 protein, which is subsequently cleaved to polypeptides of M_r 50 000 and 43 000. Occurring concomitantly is cleavage of the M_r 80 000 protein to an M_r 73 000 polypeptide. These results indicate that fully activated factor VIII may consist of subunits

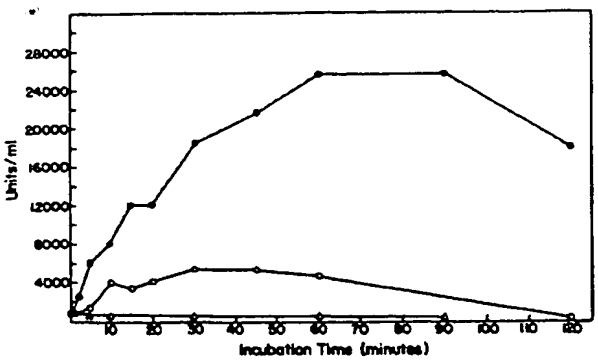


FIGURE 4: Activation of factor VIII by thrombin and the effect of $CaCl_2$ on thrombin activation. Factor VIII (110 μg/mL, ~700 units/mL) was incubated for the times shown at 37 °C with thrombin (1.5 μg/mL) in the presence of 2.5 (●), 10 (○), or 50 mM $CaCl_2$ (△). At each time point, factor VIII coagulant activity was determined, and the reaction was stopped by the addition of SDS as described under Materials and Methods.

of M_r 73 000, 50 000, and 43 000, with the latter two originating from the N-terminal portion and the M_r 73 000 from the C-terminal portion of the factor VIII precursor protein.

As shown in Figure 4, increasing the $CaCl_2$ concentration significantly alters the extent to which factor VIII is activated by thrombin. At 10 mM $CaCl_2$, only a 13-fold activation is observed, while at 50 mM $CaCl_2$ factor VIII is not activated by thrombin (Figure 5). These results are similar to the recent findings of Hultin (1985), who has shown that activation of partially pure factor VIII by thrombin in inhibited by $CaCl_2$. Comparison of the subunit structure of factor VIII activated by thrombin at 2.5, 10, and 50 mM $CaCl_2$ shows that at the higher $CaCl_2$ concentrations (10 and 50 mM) the proteolytic processing of factor VIII by thrombin is limited (Figure 5). At 10 mM $CaCl_2$, the M_r 90 000 and 80000 polypeptides are only partially cleaved to the M_r 73 000, 50 000, and 43 000 subunits, while at 50 mM $CaCl_2$ factor VIII is not cleaved at all by thrombin (Figure 5). Interestingly, at 10 mM $CaCl_2$, after maximum activity is achieved there appears to be very little change in factor VIII subunit structure yet coagulant activity greatly diminishes (Figures 4 and 5). The reasons for this decrease in activity are presently unknown. Here we show that maximum activation and proteolytic processing of factor VIII by thrombin appear to occur at physiological $CaCl_2$ concentration (2.5 mM), while partial activation or proteolysis occurs at higher $CaCl_2$ concentrations (Figures 4 and 5). In the absence of $CaCl_2$, factor VIII was proteolyzed and activated by thrombin similarly to factor VIII treated by thrombin in the presence of 2.5 mM $CaCl_2$ (data not shown).

Inactivation of Factor VIII by APC. Treatment of factor VIII with APC resulted in a dramatic decrease in coagulant activity (Figure 6). Correlated with the decrease in activity is the cleavage of M_r 90 000–210 000 proteins with the concomitant generation of an M_r 45 000 polypeptide (Figure 6, inset). The M_r 80 000 protein is not significantly proteolyzed. As previously discussed, the N-terminal amino acid sequence of the M_r 45 000 polypeptide is identical with the N-terminal sequence of the M_r 90 000–210 000 proteins (Figure 3), and the exact site at which this cleavage occurs is presently unknown. However, as proposed above for factor Xa, APC may cleave at position 336. This site precedes a very acidic region (15 Asp–Glu; 4 Lys/Arg; total of 42 amino acids) of the M_r 90 000–210 000 proteins. Significantly, cleavage of the M_r 90 000 protein at position 372, immediately following this acidic region, generates the M_r 50 000 and 43 000 subunits of thrombin–activated factor VIII. Taken together, this suggests that the acidic region between positions 336 and 372, shown in Figure 3, is of functional importance.

Activation of Factor VIII by Factor Xa. Similar to thrombin, factor Xa cleaves factor VIII at position 372 of the M_r 90 000 protein and at position 1689 of the M_r 80 000 protein (Figures 2 and 3). These cleavages would result in activation of factor VIII, as is the case for thrombin. However, factor Xa also appears to cleave factor VIII at the same site that APC proteolyzes factor VIII (Figure 2 and 3). This cleavage at position 336 would inactivate factor VIII. These results suggest that factor Xa would at best only moderately activate factor VIII and ultimately cause inactivation. Indeed, over a 2-h time course, factor Xa initially activated factor VIII only 3-fold and eventually inactivated factor VIII (Figure 7a).

During the time course treatment of factor VIII with factor Xa, the M_r 110 000–210 000 proteins were initially cleaved, with the major product being the M_r 90 000 protein (Figure 7b). This protein was cleaved to generate M_r 50 000, 45 000, and 43 000 polypeptides (Figure 7b). Subsequently, the M_r 50 000 polypeptide appeared to be cleaved to the M_r 45 000 fragment. This proteolysis correlates with the inactivation of factor VIII by factor Xa (Figure 7). Occurring concomitantly with cleavage of the M_r 90 000 protein is the cleavage of the M_r 80 000 protein to the M_r 73 000 polypeptide, which is subsequently cleaved to generate an M_r 67 000 polypeptide. This latter cleavage also correlates with factor VIII inactivation (Figure 7). Whether this cleavage, itself, is sufficient to inactivate factor VIII has not been determined.

Discussion

Recently a detailed understanding of the primary structure of factor VIII was made possible due to the isolation of factor

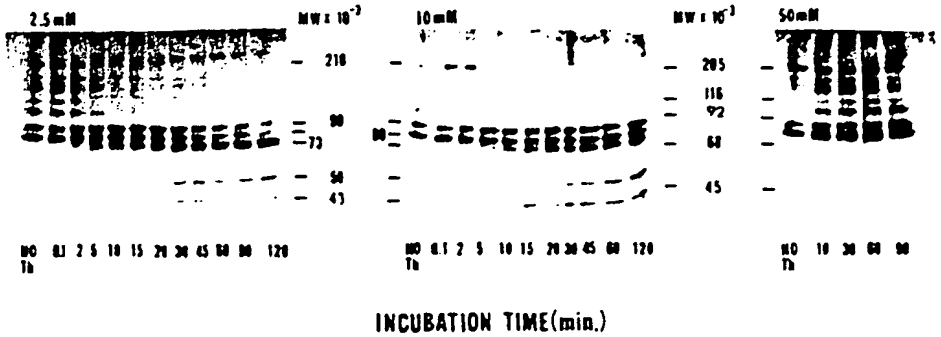


FIGURE 5: Subunit structure of thrombin-activated factor VIII. Factor VIII samples from Figure 4 that had been activated by thrombin at either 2.5, 10, or 50 mM $CaCl_2$ and subsequently made 0.5% SDS were heated to 80 °C for 5 min. The proteins were resolved on SDS–PAGE as described under Materials and Methods.

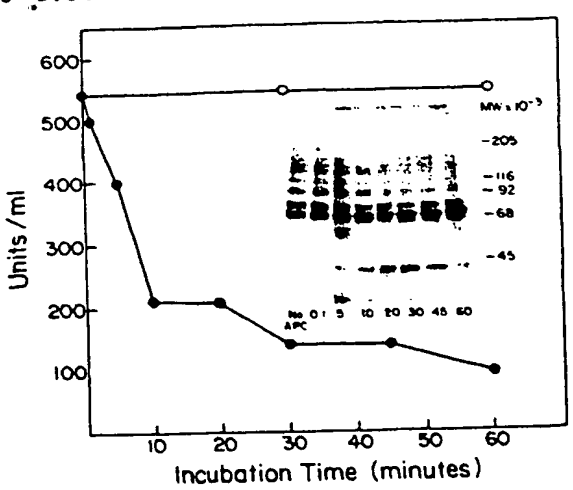


FIGURE 6: Inactivation of factor VIII by APC. To 40 μL of factor VIII (110 μg/mL) was added 5 μL of rabbit brain cephalin and subsequently incubated at 37 °C with APC (4 μg/mL) (●) or with no addition (○) for the times shown. At the end of each time point, factor VIII coagulant activity and subunit structure (inset) were determined as described in Figure 4 and under Materials and Methods.

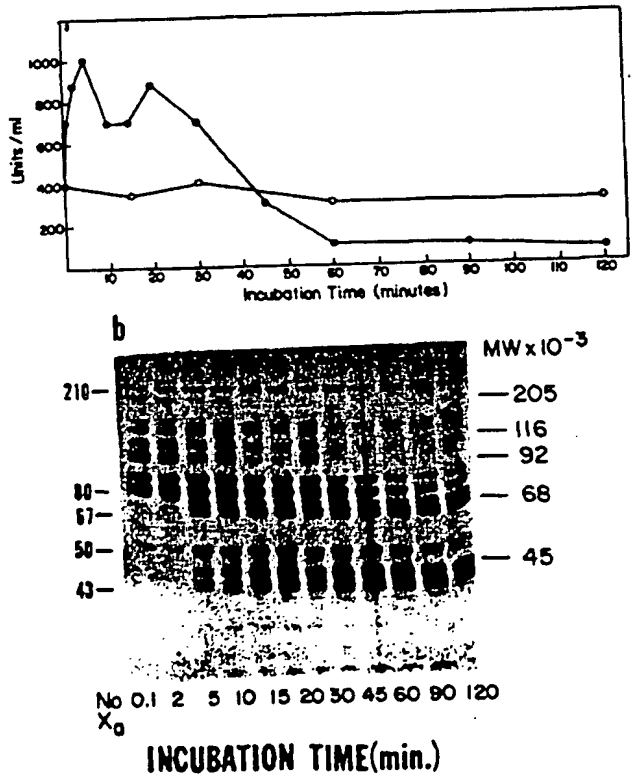


FIGURE 7: Activation of factor VIII by factor Xa. To 40 μL of factor VIII (110 μg/mL) was added 5 μL of rabbit brain cephalin and subsequently incubated with factor Xa (1 μg/mL) (●) or with no addition (○) at the times shown. At the end of each time point, factor VIII coagulant activity (a) and subunit structure (b) were determined as described in Figures 4 and 5 and under Materials and Methods.

VIII cDNA and genomic clones (Wood et al., 1984; Toole et al., 1984; Gitschier et al., 1984). The deduced amino acid sequence predicts a mature single-chain protein consisting of 2332 amino acids which, after accounting for 25 potential N-linked glycosylation sites, indicates that the single-chain form of factor VIII has an M_r >300 000. This is supported by the purification of single-chain factor VIII having M_r >300 000 from plasma (Rotblat et al., 1985). The single-chain precursor form of factor VIII appears to be readily proteolyzed in vivo and/or in vitro, yielding a species consisting of multiple subunits with M_r 80 000–210 000. As stated previously, the protease which cleaves at position 1648 to generate the M_r 80 000 protein is unknown as are the protease(s) and sites which generate the M_r 110 000–180 000 proteins seen when factor VIII is resolved by SDS–PAGE (Figures 1 and 8).

Detailed analysis of the factor VIII sequence revealed a triplicated domain structure. These domains each consist of approximately 330 amino acids and are approximately 30% homologous (Vehar et al., 1984). Interestingly, these domains also share approximately 30% homology with the triplicated domains of the plasma copper binding protein ceruloplasmin (Vehar et al., 1984). The importance of this homology as it pertains to factor VIII function is as yet not understood. The location of these domains within the factor VIII precursor is shown in Figure 8.

Thrombin activation of purified factor VIII correlates with proteolysis at positions 740, 372, and 1689 (Figure 8). These cleavages ultimately generate the M_r 73 000, 50 000, and 43 000 subunits. Cleavage at position 740 removes the C-terminal region of the M_r 110 000–210 000 proteins, generating the M_r 90 000 protein (Toole et al., 1984). Subsequently, the M_r 90 000 protein is cleaved at position 372 to generate the M_r 50 000 and 43 000 subunits. This cleavage site is between two ceruloplasmin-like domains and follows an acidic spacer region (336–372) (Figures 3 and 8). Cleavage of the M_r 80 000 protein at position 1689 to generate the M_r 73 000 subunit also follows an acidic region (positions 1649–1689) of factor VIII, which has some sequence homology with the region between positions 336 and 372 (Vehar et al., 1984).

Similar to results presented here, porcine factor VIII has been shown to be activated 70-fold by thrombin in the presence of 5 mM $CaCl_2$ (Fass et al., 1982; Loller et al., 1984). This activation was correlated with the cleavage of an M_r 82 000 protein to M_r 44 000 and 35 000 polypeptides and the cleavage of an M_r 76 000 protein to a fragment of M_r 69 000. The M_r 82 000, 44 000, and 35 000 proteins of porcine factor VIII are analogous to the M_r 90 000, 50 000, and 43 000 proteins of human factor VIII, while the M_r 76 000 and 69 000 polypeptides of porcine factor VIII are analogous to the M_r 80 000 and 73 000 proteins of human factor VIII (Fass et al., 1985). Interestingly, thrombin-activated porcine factor VIII was found to be unstable even though its subunit structure remained unchanged (Loller et al., 1984). However, it could be stabilized by factor IXa and phospholipid, suggesting that changes other than proteolysis may cause inactivation of factor VIII coagulant activity (Loller et al., 1984). This is consistent with the results of Hultin & Jesty (1981) and Rick & Hoyer (1977), who observed that inactivation of thrombin-activated factor VIII was not mediated by active thrombin. Factor VIII activated by thrombin at 10 mM $CaCl_2$ becomes inactive with time even though there is little change in subunit structure after activation (Figure 5); it is therefore possible that thrombin-activated factor VIII may not be stable at non-physiological high $CaCl_2$ concentrations. We show here that thrombin-activated factor VIII was stable for at least 1 h at 37 °C. This stability is probably a reflection of the high concentration (110 μg/mL) of factor VIII as well as the $CaCl_2$ concentration (2.5 mM) used in thrombin activation experiments.

In contrast to our results and those of Fass et al. (1982), Fulcher et al. (1983) observed that thrombin activation of

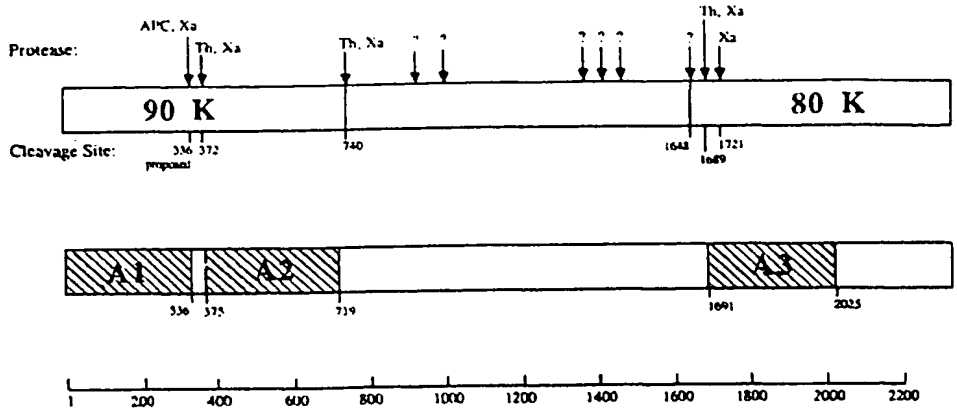


FIGURE 8: Schematic representation of proteolytic cleavage sites of thrombin (Th), factor Xa (Xa), and APC (upper diagram). The ceruloplasmin-like domains (A1, A2, and A3) and their position within the factor VIII precursor are shown in the lower diagram. The position within the factor VIII sequence is indicated by the scale.

factor VIII correlates with the generation of M_r 90 000 and 73 000 proteins. They suggest that cleavage of the M_r 90 000 protein to M_r 50 000 and 43 000 polypeptides results in inactivation of factor VIII. These investigators, however, activated factor VIII in the presence of 40 mM $CaCl_2$. As shown here and as reported by Hultin (1985), this $CaCl_2$ concentration should limit the activation of factor VIII by thrombin (Figures 4 and 5). Indeed, these authors only observed a 5-fold activation.

Inactivation of factor VIII by APC correlates with cleavage of M_r 90 000–210 000 proteins and generation of an M_r 45 000 fragment (Figure 6). This result is similar to those reported by Fulcher et al. (1984). The site at which APC cleaves has yet to be identified; however, as discussed previously, this cleavage could occur at arginine-336 (Figures 3 and 8). Since thrombin cleavage at position 372 of the M_r 90 000 protein causes activation (Figure 4), the acidic region between positions 336 and 372 may be functionally important in regulating the activity of factor VIII. Recently, Guinto & Esmon (1984) have shown that factor V proteolyzed by APC does not interact with prothrombin or with factor Xa. It is tempting to speculate that this acidic region of factor VIII may be involved in protein/protein or protein/phospholipid interactions.

Cleavage of factor VIII by factor Xa results in a modest activation followed by inactivation of factor VIII coagulant activity (Figure 6). This observation is consistent with the fact that factor Xa cleaves at the same sites as thrombin (372, 1689) as well as APC (336) (Figure 8). The inactivation of human factor VIII by factor Xa has also been observed by Triantaphyllopoulus (1979). However, purified bovine factor VIII has also been shown to be activated and stable by factor Xa at a nonphysiological low pH (Vehar & Davie, 1980). Activation of human factor VIII by factor Xa appears to correlated with the generation of M_r 73 000, 50 000, and 43 000 fragments, while inactivation appears to correlate with the cleavage of the M_r 50 000 subunit to the M_r 45 000 polypeptide. Presumably, this cleavage occurs at the same site that APC cleaves factor VIII (Figure 3). Interestingly, it has been shown that inactivation of thrombin-activated porcine factor VIII by APC correlates with the cleavage of the M_r 44 000 subunit, which is analogous to the M_r 50 000 subunit of human factor VIII (Fass et al., 1984). These results suggest that the M_r 50 000 subunit is indeed a functional subunit of activated factor VIII. Factor Xa also cleaves the M_r 73 000 subunit at position 1722 to generate an M_r 67 000 polypeptide. Whether this cleavage alters factor VIII activity has not been determined. However, this cleavage occurs within a ceruloplasmin-like domain and may release the first 20 amino acids of the A3 domain (Figure 8).

The functional as well as structural similarities of factors V and VIII have recently been recognized (Vehar et al., 1984; Fulcher et al., 1983, 1984; Church et al., 1984; Fass et al., 1985). Both of these proteins function as cofactors in the intrinsic coagulation pathway. Factor VIII in a complex with factor IXa, calcium ions, and a phospholipid surface functions in the activation of factor X. Factor V in a complex with factor Xa, calcium ions, and a phospholipid surface functions in the conversion of prothrombin to thrombin. Both proteins appear to circulate in plasma as large (M_r 300 000) single-chain precursors which are proteolytically processed to yield active cofactors (Mann et al., 1981; Rotblat et al. 1985). In both cases, the N-terminal (M_r ~90 000) and C-terminal (M_r ~80 000) portions represent the functional regions of these cofactors (Vehar et al., 1984; Toole et al., 1984; Esmon, 1979; Hibbard & Mann, 1980; Nesheim et al., 1984). Separating these regions in both factor V and factor VIII is a highly glycosylated region (M_r ~100 000). The function of this region is as yet undetermined (Vehar et al., 1984; Church et al., 1984). Amino acid sequencing also shows that factor V has homology with both factor VIII and ceruloplasmin (Church et al., 1984; Fass et al., 1985). Factor V has also been shown to contain at least a duplication of the ceruloplasmin-like domain (Fass et al., 1985). Factors VIII and V are also both activated by thrombin and factor Xa and inactivated by APC (Foster et al., 1983; Suzuki et al., 1983). Thrombin activation of either factor V or factor VIII results in the generation of polypeptides of M_r ~90 000 (from the N-termini) and ~70 000 (from the C-termini). In the case for factor VIII, the M_r 90 000 polypeptide appears to be proteolyzed further to M_r 50 000 and 43 000 polypeptides. Factor Xa proteolysis of factor V is unlike that by thrombin, and factor Xa is less efficient than thrombin in activating factor V (Foster et al., 1983). The same appears true for the activation of factor VIII by factor Xa (Figure 7). APC appears to inactivate both factor VIII and factor V by proteolyzing the N-terminal portion (molecular weight of approximately 90 000 subunit) of these factors (Figure 6; Suzuki et al., 1983). These data support the notion that factors VIII and V are regulated by thrombin, APC, and factor Xa in a very similar manner.

In this study, we have identified the cleavage sites in factor VIII of thrombin, factor Xa, and APC. This has allowed us to determine how factor VIII is qualitatively processed by these proteases. Correlation of the changes in factor VIII subunit

structure due to proteolysis by thrombin, factor Xa, or APC with changes in factor VIII coagulant activity allowed the tentative identification of the functional subunits of activated factor VIII. The results presented here suggest that the generation of the M_r 50 000, 43 000, and 73 000 subunits correlates with complete activation of factor VIII.

ACKNOWLEDGMENTS

We thank Dr. Duane Schroeder and Cutter Laboratories for supplying V_0 fractions enriched in factor VIII/vWF complexes and Dr. Walt Kisiel for the generous gift of factors IX and X and activated protein C. We greatly appreciate the support and encouragement of Dr. Richard Harkins and Dr. Richard Lawn.

Registry No. Ca, 7440-70-2; blood coagulation factor VIII, 9001-27-8; thrombin, 9002-04-4; blood coagulation factor Xa, 9002-05-5; blood coagulation factor XIVa, 42617-41-4.

REFERENCES

Bradford, M. M. (1976) *Anal. Biochem. 72*, 248–254.
Church, W. R., Jernigan, R. L., Toole, J., Hewick, R. M., Knopf, J., Knutson, G. J., Nesheim, M. E., Mann, K. G., & Fass, D. N. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 6934–6937.
Davie, E. W., Fujikawa, K., Legaz, M. E., & Kato, H. (1975) in *Proteases and Biological Control* (Reich, E., Rifkin, D. D., & Shaw, E., Eds.) Vol. 2, p 65, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Esmon, C. T. (1979) *J. Biol. Chem. 254*, 964–973.
Fass, D. N., Knutson, G. J., & Katzman (1982) *Blood 59*, 594–600.
Fass, D., Knutson, G., & Hewick, R. (1983) *Thromb. Haemostasis 50*, 262.
Fass, D., Knutson, G., Lollar, P., & Hewick, R. (1984) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 43*, 1976.
Fass, D. N., Hewick, R. M., Knutson, G. J., Nesheim, M. E., & Mann, K. G. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 1688–1691.
Fay, P. J., Chavin, S. L., Schroeder, D., Young, F. E., & Marder, V. J. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 7200–7204.
Foster, B. W., Nesheim, M. E., & Mann, K. G. (1983) *J. Biol. Chem. 258*, 13970–13977.
Fulcher, C. A., & Zimmerman, T. S. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 1648–1652.
Fulcher, C. A., Roberts, J. R., & Zimmerman, T. S. (1983) *Blood 61*, 807–811.
Fulcher, C. A., Gardiner, J. E., Griffin, J. H., & Zimmerman, T. S. (1984) *Blood 63*, 486–489.
Gitschier, J., Wood, W.I., Goralka, T. M., Wion, K. L., Chen, E. Y., Eaton, D. H., Vehar, G. A., Capon, D. J., & Lawn, R. M. (1984) *Nature (London) 312*, 326–330.
Guinto, E. R., & Esmon, C. T. (1984) *J. Biol. Chem. 259*, 13986–13992.
Hewick, R. M., Hunkapiller, M. W., Hood, L. E., & Dreyer, W. J. (1982) *J. Biol. Chem. 256*, 7990–7997.
Hibbard, L. S., & Mann, K. G. (1980) *J. Biol. Chem. 255*, 638–645.
Hultin, M. B. (1985) *Blood 66*, 53–58.
Hultin, M. B., & Jesty, J. (1981) *Blood 57*, 476–482.
Hunkapiller, M. W., Lujan, E., Ostrander, F., & Hood, L. E. (1983) *Methods Enzymol. 91*, 227–247.
Laemmli, U. K. (1970) *Nature (London) 227*, 680–685.
Loller, P., Knutson, G., & Fass, D. (1984) *Blood 63*, 1303–1308.
Mann, K. G., Nesheim, M. E., & Tracy, P. B. C. (1981) *Biochemistry 20*, 28–33.
Morrissey, J. H. (1981) *Anal. Biochem. 117*, 307–310.
Nesheim, M. E., Foster, B. W., Hewick, R., & Mann, K. G. (1984) *J. Biol. Chem. 259*, 3187–3196.
Rick, M. E., & Hoyer, L. W. (1977) *Br. J. Haematol. 36*, 585–597.
Rotblat, F., O'Brien, D. P., O'Brien, F., Goodall, A. H., & Tuddenham, E. G. D. (1985) *Biochemistry 24*, 4294–4300.
Savidge, G., Carlebjork, G., Thorell, L., Hessel, B., Holmgren, A., & Blomback, B. (1979) *Thromb. Res. 16*, 587–599.
Suzuki, K., Stenflo, J., Dahlback, B., & Teodorsson, B. (1983) *J. Biol. Chem. 258*, 1914–1920.
Toole, J. J., Knopf, J. L., Wozney, J. M., Sultzman, L. A., Buecker, J. L., Pittman, D. D., Kaufman, R. J., Brown, E., Shoemaker, C., Orr, E. C., Amphlett, G. W., Foster, W. B., Coe, M. L., Knutson, G. J., Fass, D. N., & Hewick, R. M. (1984) *Nature (London) 312*, 338–343.
Triantaphyllopoulus, D. C. (1979) *Thromb. Haemostasis 42*, 838–847.
Vehar, G. A., & Davie, E. W. (1980) *Biochemistry 19*, 401–410.
Vehar, G. A., Keyt, B., Eaton, D., Rodriguez, H., O'Brien, D. P., Rotblat, F., Oppermann, H., Keck, R., Wood, W. I., Harkins, R. N., Tuddenham, E. G. D., Lawn, R. M., & Capon, D. J. (1984) *Nature (London) 312*, 337–342.
Wood, W. I., Capon, D. J., Simonsen, C. C., Eaton, D. L., Gitschier, J., Keyt, B., Seeburg, P. H., Smith, D. H., Hollingshead, P., Wion, K. L., Delwart, E., Tuddenham, E. G. D., Vehar, G. A., & Lawn, R. M. (1984) *Nature (London) 312*, 330–337.

A SHORT POLYPEPTIDE MARKER SEQUENCE USEFUL FOR RECOMBINANT PROTEIN IDENTIFICATION AND PURIFICATION

Thomas P. Hopp, Kathryn S. Prickett, Virginia L. Price, Randell T. Libby[1], Cynthia J. March, Douglas Pat Cerretti, David L. Urdal and Paul J. Conlon

Immunex Corporation, 51 University Street, Seattle, Washington 98101. [1]Present address: Departments of Genetics and Pathology, SM-30, University of Washington, Seattle, WA 98195.

A small hydrophilic peptide of eight amino acids (AspTyrLysAspAspAspAspLys) was engineered onto the N-terminus of a variety of recombinant lymphokines for the purpose of aiding in their detection and purification from yeast supernatants or *E. coli* extracts. An antibody specific for the first four amino acids of this sequence was used as a detection reagent and for immunoaffinity purification of products under mild conditions. Because of the small size of the peptide moiety and its hydrophilic nature, the proteins were unaffected by its presence and retained a high level of biological activity. In addition, it was possible to remove the peptide via an enzymatic cleavage procedure using enterokinase.

Among the many approaches taken to improve the yield and purity of recombinant proteins, one particularly useful procedure is to express the desired polypeptide as part of a larger fusion protein[1,2]. Fusion to a protein-export signal sequence has been used to cause secretion of products from yeast[3] and *E. coli*[4] cells. Furthermore, it has recently been recognized that an attached fusion polypeptide sequence might serve as an aid to identifying or purifying the product. For example, in several cases the added polypeptide segment contains a complete second protein that binds to affinity columns via its specific substrate or ligand. These include β-galactosidase fusion proteins that bind to aminophenylthiogalactosidyl Sepharose columns[5] and protein A fusion proteins that bind to immunoglobulin columns[6]. Such fusion proteins can be highly purified in good yields in a single step by passing cell extracts or supernatants over columns of an appropriate affinity matrix, then eluting the purified fusion protein by changing conditions so that binding is no longer possible. A related approach is to use an antibody directed against the added sequence as a detection or affinity purification reagent[7-9] although the high binding affinity of most antisera and monoclonal antibodies often requires the use of denaturing conditions for elution of the product.

The fusion protein approach as several drawbacks that have not been adequately addressed in the past. First, most fusion protein products fail to fold properly into a native, active state[10]. It is possible that the added polypeptide segment is responsible for this misfolding due to unfavorable interactions during folding of the protein. This often necessitates treatment with denaturants such as 8M urea and 7M guanidine followed by refolding procedures[11,12].

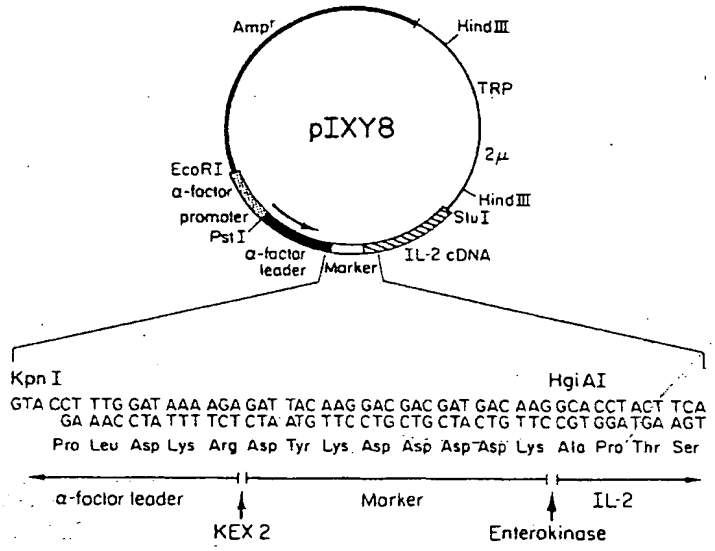
A further problem with fusion proteins is that it is difficult or impossible to remove the additional amino or carboxyl-terminal sequence from the desired product. One solution has been to use relatively harsh conditions and chemical cleavage agents such as 70% formic acid[8,12] or low pH incubations for cleavage. However, recent studies have attempted somewhat milder chemical cleavages such as hydroxylamine treatment at pH 9.0[14] or enzymatic cleavage procedures under physiological conditions. The blood clotting factor X_a has a proteolytic specificity for the tetrapeptide sequence IleGluGlyArg, and has been used to cleave human globin from a λcII protein fusion sequence[10]. A sequence allowing specific cleavage by collagenase has also been proposed[5]. Sassenfeld and Brewer[15] developed a modified ion-exchange purification technique by fusing desired proteins to a C-terminal series of arginine residues which are subsequently removed by carboxypepudase B treatment. These enzymatic processes have been successful in several instances, but often have been limited by low cleavage yields or by unwanted cleavages that occur within the desired protein sequence[10].

We decided to create a recombinant protein expression and purification system that incorporated several features of the above mentioned procedures in order to provide a fusion sequence with a combination of the most desirable properties. Here we report the development of a short amino-terminal fusion sequence AspTyrLysAsp AspAspAspLys that we refer to as a marker sequence or "Flag™" sequence for antibody mediated identification and purification of recombinant proteins. We also describe a monoclonal antibody that reacts with this sequence and can be used as an immuno-affinity purification reagent that purifies marker fusion proteins under very mild conditions. Finally, the marker sequence can be removed by treatment with the protease, enterokinase, which is specific for the four terminal amino acids of the marker sequence[16]. No harsh treatments are required at any step in this process and proteins purified by this approach retained their biological activity throughout the purification, even while the marker sequence was attached. This paper describes the expression of several such fusion proteins in *Saccharomyces cerevisiae* and *Escherichia coli*.

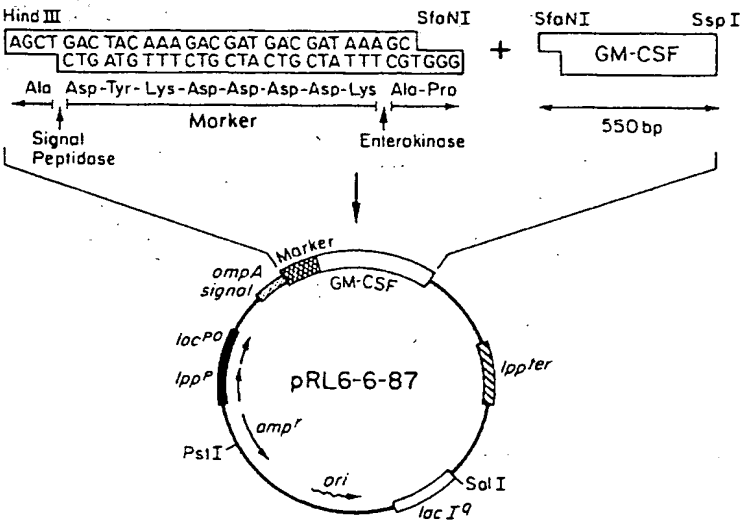
RESULTS

In order to develop this system, we performed a series of interrelated steps. The eight amino acid marker peptide was engineered onto the N-terminus of the lymphokine interleukin 2 (IL-2)[17] by means of synthetic oligonucleotides. The fusion protein was expressed in yeast and the product purified by conventional means, then used as an immunogen to produce a monoclonal antibody (4E

ific for the marker sequence. The antibody func-ed as a reagent for a number of different immuno-nical procedures including "Westerns," "dot blots," unoprecipitations, and affinity purification when led to a solid support. Furthermore, the discovery the 4E11 antibody would release its antigen when um was removed from the medium led to the devel-ent of a mild purification procedure for fusion pro-tein elution from the affinity columns. Next, treatment of the fusion proteins with enterokinase demonstrated that the enzyme was capable of removing the marker segment efficiently, with little or no observable degradation of the desired protein product. Finally, measurements of specific activity demonstrated that, for all proteins tested, no appreciable loss of activity was caused by the presence of the marker segment on the N-termini of the recombinant



Plasmid pIXY8, for expression of the IL-2 fusion in yeast. The synthetic oligonucleotides used in con-this plasmid extended from the KpnI site at the left quence shown below the plasmid diagram, to a blunt the HgiAI site near the right side, ending with the the C-terminal lys residue of the marker peptide The IL-2 coding sequence extended from a blunt end before the first codon of IL-2 (Ala) to a StuI site beyond the termination codon. The arrows below the amino acid sequence indicate the sites of cleavage by the KEX2 protease to remove the α-factor precursor sequence from the primary translation product, and by enterokinase, to remove the marker peptide from the product protein.



Plasmid pRL6-6-87, for expression of the GM-CSF protein in *E. coli*. Abbreviations: *lpp*^{p}, lipoprotein *lac*^{PO}, lactose promoter-operator region; *ompA* signal membrane protein A signal peptide sequence; GM-CSF, human granulocyte-macrophage colony stimulating factor; *lpp*^{ter}, lipoprotein terminator; *lacI*^{q}, lactose repressor; *ori*, origin of replication.

products.

Secretion of fusion proteins from yeast. Figure 1 shows the plasmid pIXY8, used for production of the IL-2 fusion protein in yeast. Similar constructs with the ADH2 promoter replacing the α-factor promoter were used to express granulocyte colony stimulating factor (G-CSF)[18], interleukin 3 (IL-3)[19], interleukin 4 (IL-4)[20] and granulocyte-macrophage colony stimulating factor (GM-CSF)[21] fusion proteins. All of these proteins were secreted into culture media by yeast. Each had the expected molecular weight for the correctly processed form (i.e. with leader peptide removed) and yielded the expected sequence of the marker peptide on N-terminal amino acid sequence analysis. The IL-2 fusion protein was purified by HPLC for use as an immunogen, while the other fusion proteins were purified by the 4E11 antibody immunoaffinity chromatography procedure (below). For comparison, essentially identical vectors were prepared that expressed each protein without the marker peptide. These products were recovered from the yeast culture media and purified to homogeneity by conventional techniques including HPLC and ion exchange chromatography[22].

Expression of GM-CSF fusion protein in *E. coli*. The construction of the plasmid pRL6-6-87 for expression and secretion of the GM-CSF fusion protein in *E. coli* is outlined in Figure 2. This plasmid allows the secretion of the marker peptide GM-CSF fusion protein by means of the signal peptide from the outer membrane protein OmpA. The product obtained from *E. coli* cultures had the expected molecular weight for the marker peptide GM-CSF fusion and yielded an N-terminal amino acid sequence corresponding to the marker peptide sequence.

The marker-specific antibody. The isotype of the 4E11 antibody is IgG 2B. It was found to be reactive with proteins bearing the marker peptide sequence in a variety of procedures, including ELISAs, dot blots, Western blots, immunoprecipitation and affinity chromatography, as described below. The antibody was found to react with all of the marker peptide fusion proteins that we have produced. The antibody exhibits no reactivity [illegible] marker products, or any component pres[illegible] extracts, or in yeast culture medium.

Purification. Figure 3 shows the resul[illegible] affinity purification chromatograms on aff[illegible] made with the 4E11 antibody. In Figure 3[illegible] supernatant obtained by fermenting yeast [illegible] GM-CSF expression vector was passed over the [illegible] purify the fusion protein that had been secret[illegible] medium. Medium components were removed [illegible] with PBS containing 0.5 mM $CaCl_2$, while the [illegible] remained bound to the antibody. Subsequently, [illegible] PBS containing EDTA dissociated the mark[illegible] antibody complex and released the GM-CSF [illegible] tein as a purified product. The multiple molecular [illegible] species eluting from the column are typical of [illegible] teins secreted from yeast and result from heter[illegible] glycosylation by yeast cells. All bands were [illegible] GM-CSF, based on Western analyses using 4E[illegible] anti. GM-CSF monoclonal antibody as develo[illegible] agents.

Figure 3B shows the results of affinity chromat[illegible] of an extract of *E. coli* cells that had been trans[illegible] with pRL6-6-87 in order to produce the GM-[illegible] protein. Chromatography was carried out as [illegible] yeast GM-CSF fusion, except that 1 mM Ca[illegible] during washing and 0.1 M glycine HCl pH 3.0 [illegible] used to elute the product. The GM-CSF fusion [illegible] eluted as a single molecular weight species beca[illegible] does not glycosylate proteins. The product was [illegible] pure after this single chromatographic step.

The binding of the 4E11 antibody to the [illegible] peptide is dependent on the presence of calcium [illegible] property has been reported for a few other antib[illegible] the past[23]. We observed that if insufficient [illegible] calcium were present in washing buffer, then the [illegible] proteins would leak from the affinity column [illegible] though they had bound quantitatively when yeast [illegible] natant or *E. coli* extract was passing over the colu[illegible]

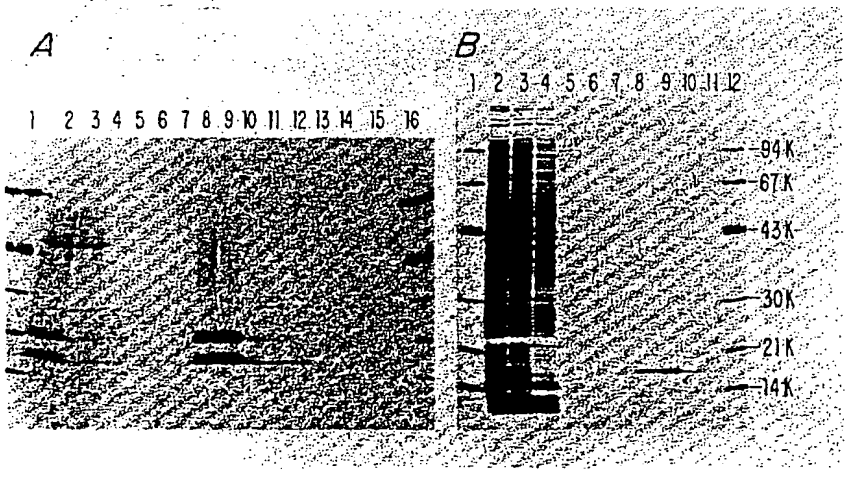


FIGURE 3 Affinity purification of recombinant fusion proteins. Panel A: Silver-stained polyacrylamide gel of yeast GM-CSF fusion protein purification. Lanes are: 1: molecular weight standards; 2: yeast supernatant; 3: flow through material; 4–6: sequential PBS/0.5 mM $CaCl_2$ washes; 7–15: sequential PBS/2.0 mM EDTA elutions; 16: molecular weight standards. Panel B: Silver-stained gel of *E. coli* GM-CSF fusion protein purification. Lanes are: 1: molecular weight standards; 2: *coli* supernatant; 3: flow through material; 4–6: sequent[illegible] PBS/0.5 mM $CaCl_2$ washes; 7–11: sequential 0.1 M Gly-H[illegible] pH 3 elutions; 12: molecular weight standards. Numbers right indicate M, values for the standard proteins (in kiloDa[illegible] tons). The 94K standard was omitted from the gel shown panel A.

ing the effect of various concentrations of $CaCl_2$ in ning buffer, we determined that concentrations of ium above 0.3 mM were necessary to retain the fusion teins on the column. Given this calcium dependence, it found that rapid elution of fusion proteins could be ved by using EDTA in the elution buffer. It was also ible to elute proteins simply by using a calcium free ion buffer after the columns had been washed with containing calcium. However, under these conditions luted fusion protein tended to spread through more ons than when the elution buffer contained EDTA.

terokinase treatment. Figure 4 shows the results of okinase treatment of the IL-2 fusion protein. In- ing amounts of enterokinase were added to identical ots of the IL-2 fusion protein, then the samples were ted for 16 hr at 37°. As the concentration of enzyme creased, a component appeared at the molecular of the authentic IL-2. At the highest concentra- of enzyme, the conversion of fusion protein to ein protein was complete, and an approximately l amount of authentic product had been formed. n blot analyses using 4E11 and anti-IL-2 monoclo- ibodies confirmed the identities of the products the silver gel. Only the higher molecular weight was reactive with 4E11, but as expected both the and lower molecular weight species were reactive anti-IL-2 antibody. Amino terminal sequence on the resulting cleavage products indicated that usion protein was cleaved after the second lysine arker peptide yielding the expected sequence for enic IL-2 N-terminus. As can be seen in Figure 4, ge is complete, and in this case results in only ounts of detectable lower molecular weight by- formed by enterokinase cleavage within the IL-2 Essentially the same cleavage pattern was ob- h the other fusions as well. In no case were ion products present in quantities greater than ual protein (determined by sequence analysis).

ion of biological activity. The marker peptide oteins expressed in yeast and *E. coli* were all ly active despite the presence of the marker as is seen in Table 1. In all cases the levels of activity obtained with the fusion proteins were le to wild-type recombinant proteins expressed e marker sequence. In the case of GM-CSF the ivity values in Table 1 were obtained before removal of the marker sequence by enteroki- nearly identical specific activities obtained indi- me yield of cleaved product is probably near

ION

arker peptide fusion system described in this prises a unique and widely useful technique n identification and purification. In addition, vations have shown that the marker peptide is with heterologous expression systems: one, *coli* OmpA signal and the yeast pre-pro alpha tor sequences, when fused to the marker pep- rrectly processed by their respective proteases retion of marker peptide-protein fusions with N-terminus. And second, because many inves- ve reported problems in the N-terminal proc- icrobially expressed mammalian proteins[24–26], of the marker peptide to protect the N- the desired product may be another impor- of this system.

ibility to produce authentic N-termini upon treatment is an improvement over a number ion protein approaches. Those that require chemical cleavages using Asp-Pro[12] or Asn-Gly[13] directed reagents, for example, must necessarily leave a proline or glycine at the N-terminus of the product. Although we have not yet tested the ability of enterokinase to cleave the marker sequence from N-termini containing all of the 20 possible amino acids, we have found that it is capable of cleaving products with N-terminal Glu, Ala, Thr, Leu and Ile residues. This suggests that this procedure will be useful for a wide range of N-termini including charged and uncharged, hydrophobic and hydrophilic residues.

There are several requirements that should be met by an efficient detection and purification system based upon fusion polypeptide expression: First, the added marker segment should not interfere with the native folding of proteins to which it is attached. Second, the marker peptide sequence should be intrinsically water soluble and should retain a high degree of exposure in the aqueous environment of the protein, so that it can readily interact with the affinity purification substrate. Third, it should be useful in an affinity purification step that requires only very mild media, and be elutable with a non-denaturing and inexpensive eluant. Finally, the marker peptide should be easy to remove and the product protein should not have any amino acids added or deleted once the marker peptide has been removed. The Flag™ peptide fusion system was designed to possess all of these properties, and our data with several recombinant proteins suggest that it may prove to be a universal purification system for proteins expressed in heterologous organisms.

Several factors were considered in choosing the specific sequence of the marker peptide moiety. We chose to limit the marker peptide sequence to only eight amino acids because it can easily be encoded in a single synthetic oligonucleotide, and because the longest trypsinogen pro-sequences are of this length. We therefore could be reasonably sure that the trypsin-activating enzyme, enterokinase, would work efficiently to release the peptide. Additionally, because antibodies require up to six or seven amino acids for avid binding interactions, we reasoned that eight amino acids should be the minimum sequence capable of strong binding to an antibody while allowing one or more of the last amino acids on the C-terminal end to act as a spacer to separate the antibody binding portion from interference with the bulk of the protein. Finally, the five C-terminal amino acids of the marker sequence represent the minimal enterokinase specificity site, AspAspAspAspLys.

The choice of Lys at position three of the marker

TABLE 1 Expression level and specific activity of marker fusion proteins.

Protein	mg/l[a]	Specific Activity (U/mg × 10^{-6}) With Marker	Specific Activity (U/mg × 10^{-6}) Without Marker[b]
	Expressed in Yeast		
IL-2	1–3	240 ± 20	280 ± 40
G-CSF	10–20	65 ± 10	12 ± 1.0
IL-3	15–20	49 ± 19	42 ± 12
IL-4		5.4 ± 1.6	N.D.[c]
	Expressed in *E. coli*		
GM-CSF	15–20	150 ± 60	140 ± 60

[a]Determined by dot blot assay using 4E11 antibody to detect marker peptide containing material.
[b]Proteins in this column (except *E. coli* GM-CSF) were produced without the marker segment and purified by conventional means.
[c]Not determined.
[d]In this case, specific activities were determined on the same sample before and after enterokinase treatment to remove the marker segment.

sequence causes the marker peptide to contain the hexapeptide sequence, LysAspAspAspAspLys, that has a maximum value on the hydrophilicity scale of Hopp and Woods[27]. Such maximally hydrophilic sequences have been proven to express strong antigenicity and are correspondingly likely to adopt a highly exposed conformation in the three dimensional folding of a protein[28]. As can be seen in Figure 5, it is impossible for any other region of a protein to have a higher hydrophilicity value than this maximally hydrophilic sequence, so the marker segment is virtually guaranteed to be exposed at the surface of any fusion protein. Therefore it can always be expected to be available for binding to antibody. Perhaps most importantly, the strong predilection for externalization should guarantee that the marker segment will not interfere in the adoption of a native conformation by the remainder of the protein.

In addition to the hydrophilic effects of Lys at position three, several other considerations influenced the choice of amino acids at the N-terminus of the marker peptide. Aromatic amino acids have been recognized as major factors in antigen-antibody interactions[29] so a tyrosine was placed at position 2, flanked by charged amino acids. Recent evidence suggests that aromatic residues that are flanked by charged sequences are more likely to be involved in antigenic sites than are other aromatic residues in less polar environments[28]. The decision to place an Asp residue at the N-terminus was made in part because the negative charge on the Asp should aid in exposing the Tyr to antibody, as mentioned above, and in part because, with the inclusion of the Asp at position 1, a total of eight charges are to be found on the marker peptide moiety, including the N-terminal amino group. This preponderance of charged residues was expected to make it likely that antibody binding would be heavily dependent on charge-charge interactions, and therefore might be highly susceptible to elution with commonly used salt solutions such as 2 M $MgCl_2$ or 1 M NaCl[9]. In the end, the serendipitous discovery that Ca^{+2} was involved in the charge-dependent binding of the marker sequence to the 4E11 antibody made even these mild salt treatments unnecessary.

We have seen that fusion proteins retain the appropriate specific activity even with the marker segment still attached, and that this activity can be maintained after enterokinase treatment to remove the marker sequence. Comparisons of several of these fusion protein products with their natural counterparts (Table 1) demonstrated that the presence of the marker did not decrease the specific activity of the fusion proteins relative to the same proteins with no extraneous amino acids added. We have recently begun using a larger version of the 4E11 column to prepare proteins in milligram quantities. This level of scale-up required no special procedures or equipment, and can still be done as a bench top experiment. Further scale up for production of gram or kilogram quantities is contemplated, and will be limited only by availability of antibody and enterokinase. One advantage of the 4E11 antibody is that it can be purified on an affinity column comprised of chemically synthesized marker peptide attached to a solid support and eluted with EDTA.

One area that remains potentially problematic is the provision of adequate enterokinase for this process, both in terms of quantity and quality. We sampled commercial sources of enterokinase, but found that our fusion proteins were digested into small fragments, presumably by contaminating chymotrypsin, trypsin and elastase that are likely to be present in these partially purified preparations. Our own crude bovine intestinal preparations also caused substantial unwanted hydrolysis, until we used the

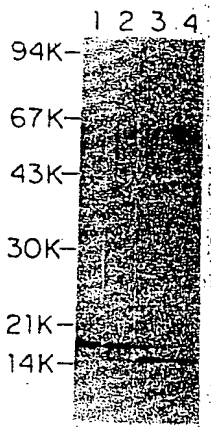


FIGURE 4 Enterokinase digestions. The IL-2 fusion p... incubated with increasing amounts of bovine ... and the digestion was analyzed by silver-staining PA... are: 1: purified IL-2 fusion protein (200 ng) no... fusion protein + 2 ng enterokinase; 3: fusion protein ... enterokinase; 4: fusion protein + 20 ng enterokin...

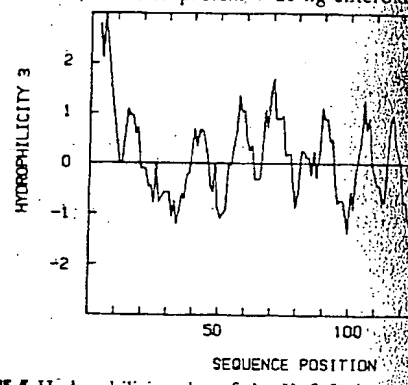


FIGURE 5 Hydrophilicity plot of the IL-2 fusion prote... profile was generated using the updated HYDRO... of Hopp[19]. The scale is oriented so that hydrophilic ... hydrophobic is at bottom. Valley regions are expected ... buried portions of the polypeptide, whereas peaks are ... ed to be exposed at the surface of the protein. The pro... peak near the N-terminus results from the extremely ... philic hexapeptide, LysAspAspAspAspLys, contained ... the marker peptide sequence.

protocol of Liepnieks and Light[30] for removing ... other proteases. The fact that a minor amount of u... ed cleavage is seen in some cases when the marker ... is removed (Fig. 4) suggests that traces of contamin... proteolytic activity may still be present in our ... preparations of enterokinase. An ideal solution would ... to clone enterokinase and express it in a recombi... organism. This would provide starting material ... much lower levels of contaminating proteases, and ... higher levels of enterokinase. Furthermore, with ... enterokinase gene in hand, it might be possible to ... neer a smaller form of the molecule, lacking the hy... phobic portion that binds it to the membranes of ... intestinal villi. This would simplify purification and ... crease the mass of enterokinase needed for cleavage of ... marker segment.

Despite the need for further development of enterok... nase, this system for fusion protein detection and purifi... tion already represents a useful technique. It offers ... possibility of using a single procedure for the purificati...

multiple fusion proteins. Although it is also possible to ify fusion proteins from yeast medium or *E. coli* cells conventional means such an ion exchange or reverse e chromatography, these procedures require new hod development for each new protein, whereas with ker fusions, the same process is applicable to all eins. Finally, because the marker segment does not ar to have decreased the biological activity of any of roteins that it has been placed on, it may not always ecessary to remove the marker segment in order to an active product. In such cases, this useful "han- can be retained on the molecule, enabling investiga- o readily detect and manipulate their recombinant in products.

RIMENTAL PROTOCOL

n plasmid construction. The yeast vector used for protein sion has been described previously[31]. This vector contains ces from pBR322 that allow selection (Amp^r) and replica- *E. coli*, as well as the yeast *Trp1* gene and 2μ origin of ion for selection and autonomous replication in a trp1 rain. Expression of foreign genes is under control of the promoter[5] or the ADH2 promoter[32] and secretion is by the α-factor leader peptide. To generate the IL-2 ion vector pIXY8, the mature coding region of IL-2[17] was -frame to the marker peptide and the α-factor leader by a synthetic oligonucleotide linker encoding the five C- amino acids of the α-factor leader and the eight amino he marker peptide (Fig. 1). The vectors that directed the ion of the other products were generated by two modifi- pIXY8. First, the α-factor promoter was removed by ng the plasmid with EcoRI and PstI, then inserting the romoter using a synthetic oligonucleotide linker. Sec- marker and appropriate protein coding sequences were place of the IL-2 sequence (Fig. 1) and linked with a oligonucleotide that extended to the HpaI site.

n of yeast strains. *S. cerevisiae* strain XV218/(a/α-trp-1) n an either selective medium [YNB -trp, consisting of ast Nitrogen Base (Difco), 0.5% Casamino acids, 2% 10 μg/ml adenine and 20 μg/ml uracil] or rich medium onsisting of 1% Yeast Extract, 2% peptone and 1% supplemented with 80 μg/ml adenine and 80 μg/ml . Yeast transformations were done by selecting for Trp^+ nants[33]. Cultures were grown for biological assay by ng 20–50 ml of rich medium with the appropriate strain ng the cultures at 30°C to stationary phase. Cells were ved by centrifugation and the medium was filtered 0.45μ cellulose acetate filter. Sterile supernates were °C. Larger scale fermentations were done in a 10 liter wick Microferm fermentor. Cells were removed from um using a Millipore Pellicon filtration system.

uction of the *E. coli* vector. Plasmid pIN-III-$OmpA_3$ is expression vector regulated by the tandem *lpp*^{p} (lipopro- moter)/*lac*^{po} (lactose-promoter-operator) that contains a *ompA* leader sequence for protein secretion[34]. Construc- the GM-CSF expression vector was accomplished by digestion at the unique BamHI site of pIN-III- ollowed by its conversion to blunt ends by treatment e transcriptase (Boehringer-Mannheim). The vector quently restricted with HindIII, and used in a three- n with a synthetic oligonucleotide encoding the marker quence and cDNA encoding GM-CSF to produce as outlined in Figure 2.

of *E. coli*. Plasmid pRL6-6-87 was introduced by ation into *E. coli* strain JM107, (Δ*lac*, *pro*, *thi*, *strA*, *endA*, *sbRA/F'*, *traD36*, *proAB*$^+$, *lacI*^{q}-ZΔM15) which was grown M9 minimal medium containing 1% (w/v) methionine m (DIFCO) and ampicillin (50 μg/ml) to an OD_{600} of es were induced following addition of isopropyl-β-D- pyranoside (IPTG) and cyclic 3'-5' adenosine mono- (cAMP) to 2mM and 4mM, respectively, and allowed ate the GM-CSF fusion protein for 2–4 hr. Cells were by centrifugation, and pellets either stored at −70°C or rocessed for extraction and purification of marker eins. *E. coli* pellets were extracted by the following the pellet from 500 ml of culture was suspended in 50 mM NaCl, 50 mM NaH_2PO_4, pH8.4, to which 1 mM nylsulfonylfluoride had been added immediately prior After freezing (−70°C) and thawing three times to lyse ells the sample was incubated at 37°C for 30 min to complete lysis and to extract the protein product. The viscous extract was treated by dounce homogenization to achieve a uniform solution, then centrifuged at 25,000 g for 45 min at 4°C. The supernatant was adjusted to 0.5 mM $CaCl_2$, recentrifuged if necessary to remove any resulting precipitate, then applied to the affinity column.

Preparation of immunogens. Palmitic acid conjugated peptides were produced by solid phase chemical synthesis as described previously[30]. The antigenic marker peptide had the sequence AspTyrLysAspAspAspAspLysGlyProLysLysGly to which palmitoyl moieties had been attached on the epsilon amino groups of the two C-terminal lysines. It is referred to as CDP-marker (C-terminal dipalmitoyl marker peptide). A second palmitoyl peptide, NDP-GM1 (N-terminal dipalmitoyl GM-CSF peptide 1) was used as a non-specific binding control. It had the structure LysGlyGlyGluSerPheLysGluAsnLeuLysAspPheLeuValGly, and also possessed two palmitoyl moieties, in this case attached to the two amino groups of the N-terminal lysine residue. For purification of the IL-2 marker peptide fusion protein, supernatants of yeast expressing the IL-2 fusion protein were applied to a reverse phase HPLC column. The IL-2 fusion product was eluted from the column using a gradient of acetonitrile, as described previously[22].

Immunization. BALB/c female mice were purchased from the Jackson Laboratories (Bar Harbor, ME) and maintained in our animal facility. Mice were immunized subcutaneously with 250 μg of IL-2 fusion protein emulsified in Freund's complete adjuvant, followed with 125 μg of the same protein emulsified in Freund's incomplete adjuvant four weeks later. Two weeks after the second inoculation, a serum antibody titer to IL-2 and the marker peptide was measured by "dot-blot". The animal was then challenged with 10 μg of protein intravenously four days prior to fusion.

Hybridoma derivation. Four days after the intravenous boost, the animals were sacrificed, their spleens removed, and a single cell suspension prepared. The splenocytes were fused to the HAT sensitive myeloma cell, NS-1. The resulting hybridomas were then assayed for the production of antibodies to the marker peptide seven to ten days later by ELISA (see below). One hybridoma antibody consistently produced a positive reaction specifically with the marker peptide moiety. This cell line, designated 4E11, was then cloned by limiting dilution, isotyped, and further characterized.

ELISA. Various peptide solutions (CDP-marker, or NDP-GM1) were applied to HA plates (Millipore, Bedford, MA) at a concentration of 40 ng per well and allowed to incubate for 30 min at room temperature. Nonspecific protein binding sites were blocked by an incubation with 3% bovine serum albumin in Tris buffered saline, pH 7.0 (TBSA) for 1 hr at room temperature. Hybridoma supernatants were added and the plates incubated for 1 hr. Following this incubation, the plates were washed with PBS and an alkaline phosphatase labeled goat anti-mouse antibody (Sigma Chemical, St. Louis, MO) was added. Following a 1 hr incubation, the plates were washed several times with PBS and a colorimetric indicating reagent was added (substrate tablets, Sigma Chemical). Contents of each HA plate were then transferred to a polystyrene 96 well plate (Linbro/Titertek, Flow Laboratories, McLean, VA) and the absorbance at 405 nm determined on a Titerscan (Flow Laboratories).

Production and purification of 4E11 antibody. Pristane-primed BALB/c mice were injected IP with 1×10^6 hybridoma cells. Ten to twenty days later, the ascitic fluid was recovered, centrifuged at 1000 × *g* for 30 minutes at 4°C, passed through cotton gauze, and the supernatant stored at −20°C until needed. The monoclonal antibody 4E11 was purified from ascites fluid using MAPS II Protein A Affigel (Bio-Rad, Richmond, CA) affinity chromatography. The purified antibody was found to be homogeneous by SDS-PAGE analysis.

4E11 column preparation. Purified 4E11 immunoglobulin was concentrated by ultrafiltration. After dialysis against 0.1M Hepes buffer, pH 7.5 at 4°C the antibody was coupled to Affigel-10 (Bio-Rad) in accordance with the manufacturer's instructions. A typical antibody-coupled gel contained from 1.5 to 4.5 mg antibody/ml of gel. Columns of 4E11 coupled gel of 1.5 ml bed volume were prepared in polypropylene columns (Bio-Rad) and washed with 15 ml PBS, 15 ml 0.1M glycine HCl, pH 3.0, and stored at 4°C in PBS/0.02% sodium azide.

4E11 column chromatography. Yeast culture filtrates were brought to physiological levels of salt and pH by adding 10X PBS, and made 0.5 mM in $CaCl_2$ by adding 1 M $CaCl_2$, and then loaded onto the 1.5 ml column of 4E11 coupled Affigel 10 under gravity flow. *E. coli* extracts did not require any further additions because

the extraction medium contained physiological levels of salt and pH as well as 0.5 mM $CaCl_2$. Up to 100 ml of filtrate were passed over the column, depending on the level of expression of the recombinant protein. After loading, the column was washed with three to five aliquots of 3 ml of PBS containing 0.5 mM $CaCl_2$. Elution was carried out with PBS lacking $CaCl_2$ and containing 2.0 mM Na_2 EDTA or with 0.1M glycine HCl pH 3.0. Each elution fraction was 1 ml. Yields of purified proteins were determined by amino acid analysis, and were typically 15–40% of the theoretical maximum assuming a 2:1 antigen to antibody binding ratio.

Enterokinase treatment. Enterokinase was purified from bovine intestine by the procedure of Liepnieks and Light[30]. Samples were also provided by A. Light of Purdue University. For enterokinase treatment, fusion proteins eluted from the antibody column were made 10 mM in Tris-HCl (pH 8) and adjusted to pH 8.0 by addition of 1N NaOH. For certain samples, the reaction mixture was made 40 mM in octyl-β-D-glucoside. Following the addition of an appropriate amount of bovine enterokinase (1–10% by weight; typically 0.2–2% by molarity), the reaction mixture was incubated for 16 hours at 37°C. Enterokinase dilutions were made from a 1 mg/ml stock solution of enzyme in 10 mM Tris-HCl, pH 8 kept at −70°C.

Bioassays. The activity of IL-2 was measured using the murine IL-2 dependent T-cell line CTLL-2[36]. The activity of GM-CSF was measured in a human bone marrow proliferation assay[37] and the activity of IL-3 was measured by FDC-P2 cell proliferation[38]. IL-4 and G-CSF were assayed as described[39,40]. Specific activities were derived by measuring the biological activities of purified samples of each protein, after quantifying by amino acid analysis.

Acknowledgments

The authors wish to thank David Amberg, Janet Merriam, Aline Akelis, Norman Boiani, Bruce Hess, and David McMasters for excellent technical assistance. We would also like to thank Linda Troup for preparation of the manuscript.

Received 19 April 1988; accepted 5 July 1988.

References

1. Itakura, K., Hirose, T., Crea, R., Riggs, A. D., Heyneker, H. L., Bolivar, F., and Boyer, H. W. 1977. Expression in *Escherichia coli* of a chemically synthesized gene for the hormone somatostatin. Science 198:1056–1063.
2. Goeddel, D. V., Kleid, D. G., Bolivar, F., Heyneker, H. L., Yansura, D. G., Crea, R., Hirose, T., Kraszewski, A., Itakura, K., and Riggs, A. D. 1979. Expression in *Escherichia coli* of chemically synthesized genes for human insulin. Proc. Natl. Acad. Sci. USA 76:106–110.
3. Brake, A. J., Merryweather, J. P., Coit, D. G., Heberlein, U. A., Masiarz, F. R., Mullenback, G. T., Urdea, M. S., Valenzuela, P., and Barr, P. J. 1984. α-Factor-directed synthesis and secretion of foreign proteins in *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA 81:4642–4646.
4. Villa-Komaroff, L., Efstratiadis, A., Broome, S., Lomedico, P., Tizard, R., Naber, S. P., Chick, W. L., and Gilbert, W. 1978. A bacterial clone synthesizing proinsulin. Proc. Natl. Acad. Sci. USA 75:3727–3731.
5. Germino, J. and Bastia, D. 1984. Rapid purification of a cloned gene product by genetic fusion and site-specific proteolysis. Proc. Natl. Acad. Sci. USA 81:4692–4696.
6. Abrahmsen, L., Moks, T., Nilsson, B., and Uhlen, M. 1986. Secretion of heterologous gene products to the culture medium of *Escherichia coli*. Nucleic Acids Res. 14:7487–7500.
7. Munro, S. and Pelham, H. R. B. 1984. Use of peptide tagging to detect proteins expressed from cloned genes. EMBO J. 3:3087–3093.
8. Ikehara, M., Ohtsuka, E., Tokunaga, T., Nishikawa, S., Uesugi, S., Tanaka, T., Aoyama, Y., Kikyodani, S., Fujimoto, K., Yanase, K., Fuchimura, K., and Morioka, H. 1986. Inquiries into the structure-function relationship of ribonuclease T1 using chemically synthesized coding sequences. Biochemistry 83:4695–4699.
9. Wilchek, M., Miron, T., and Kohn, J. 1984. Affinity chromatography. Methods in Enzymol. 104:3–55.
10. Nagai, K. and Thogerson, H. C. 1987. Synthesis and sequence specific proteolysis of hybrid proteins produced in *Escherichia coli*. Methods Enzymol. 153:461–481.
11. Schoner, R. G., Ellis, L. F., and Schoner, B. E. 1985. Isolation and purification of protein granules from *Escherichia coli* cells overproducing bovine growth hormone. Bio/Technology 3:151–154.
12. Sung, W. L., Yao, F. L., Zahab, D. M., and Narang, S. A. 1986. Short synthetic oligodeoxyribonucleotide leader sequences enhance accumulation of human proinsulin synthesized in *Escherichia coli*. Proc. Natl. Acad. Sci. USA 83:561–565.
13. Nilsson, B., Holmgren, E., Josephson, S., Gatenbeck, S., Philipson, L., and Uhlen, M. 1985. Efficient secretion and purification of human insulin-like growth factor I with a gene fusion vector in *Staphylococci*. Nucleic Acids Res. 13:1151–1162.
14. Moks, T., Abrahmsen, L., Osterlof, B., Josephson, S., Ostling, M., Enfors, S., Persson, I., Nilsson, B., and Uhlen, M. 1987. Large-scale affinity purification of human insulin-like growth factor I from culture medium of *Escherichia coli*. Bio/Techno
15. Sassenfeld, H. M. and Brewer, S. J. 1984. designed for the purification of recombinant p gy 2:76–81.
16. Maroux, S., Baratti, J., and Desnuelle, P. specificity of procine enterokinase. J. Biol. Che
17. Taniguchi, T., Matsui, H., Fujita, T., Takao Yoshimoto, R., and Hamuro, J. 1983. Structur cloned cDNA for human interleukin-2. Nature
18. Nagata, S., Tsuchiya, M., Asano, S., Kazira moto, O., Hirata, Y., Kubota, N., Obeda, M., N M. 1986. Molecular cloning and expression granulocyte colony stimulating factor. Nature 3
19. Yang, Y-C., Ciarletta, A. B., Temple, P. A., Chu Witek-Giannotti, J. S., Leary, A. C., Kriz, R., D G. G., and Clark, S. C. 1986. Human IL-3 (mult by expression cloning of a novel hematopoietic to murine IL-3. Cell 47:3–10.
20. Yokota, T., Otsuka, T., Mosmann, T., Banchere Blanchard, D., DeVries, J. E., Lee, F., and Arai, and characterization of a human cDNA clone, h B-cell stimulatory factor 1, that expresses B-cell a activities. Proc. Natl. Acad. Sci. USA 83:5894–58
21. Cantrell, M. A., Anderson, D., Cerretti, D. P., Pri K., Tushinski, R. J., Mochizuki, D. Y., Larsen, A., S., and Cosman, D. 1985. Cloning, sequence, human granulocyte/macrophage colony stimulati Acad. Sci. USA 82:6250–6254.
22. Urdal, D. L., Mochizuki, D., Conlon, P. J., March M. L., Eisenman, J., Ramthun, C., and Gillis, S. purification by reversed-phase high performance phy. J. Chromatography 296:171–179.
23. Maurer, P. H., Clark, L. G., and Liberti, P. A. 19 polypeptides (polyamino acids): Calcium-depen dent antibodies. J. Immunol. 105:567–573.
24. Hitzeman, R. A., Leung, D. W., Perry, L. J., K H. L., and Goeddel, D. V. 1983. Secretion of hu yeast. Science 219:620–625.
25. Aston, R., Cooper, L., Holder, A., Ivanyi, J., an Monoclonal antibodies to human growth hormo between pituitary and genetically engineered form 22:271–275.
26. Kronheim, S. R., Cantrell, M. A., Deeley, M. C., Ma P. J., Anderson, D. M., Hemenway, T., Merriam, and Hopp, T. P. 1986. Purification and characte interleukin-1 expressed in *Escherichia coli*. Bio/Te 1082.
27. Hopp, T. P. and Woods, K. R. 1981. Prediction of determinants from amino acid sequences. Proc. Na 78:3824–3828.
28. Hopp, T. P. 1986. Protein surface analysis: Metho antigenic determinants and other interaction sites. J ods 88:1–18.
29. Sela, M. and Arnon, R. 1960. Studies on the che antigenicity of proteins. Biochem. J. 75:91–102.
30. Liepnieks, J. J. and Light, A. 1979. The preparation bovine enterokinase. J. Biol. Chem. 254:1677–1683
31. Price, V., Mochizuki, D., March, C. J., Cosman, I Klinke, R., Clevenger, W., Gillis, S., Baker, P., and Expression, purification and characterization of rec granulocyte-macrophage colony-stimulating factor leukin-2 from yeast. Gene 55:287–295.
32. Beier, D. R. and Young, E. T. 1982. Characterizatio region upstream of the ADR2 locus of *S. cerevisiae*. 728.
33. Hinnen, A., Hicks, J. B., and Fink, G. R. 1978. Tr yeast. Proc. Natl. Acad. Sci. USA 75:1929–1933.
34. Libby, R. T., Braedt, G., Kronheim, S. R., March, C. Chiaverotti, T. A., Tushinski, R. J., Mochizuki, D. and Cosman, D. 1987. Expression and purification granulocyte-macrophage colony stimulating factor fr *coli* vector. DNA 6:221–229.
35. Hopp, T. P. 1984. Immunogenicity of a synthetic Enhancement by conjugation to a fatty acid carrier. M 21:13–16.
36. Gillis, S., Ferm, M. M., Ou, W., and Smith, K. 197 factor: parameters of production and a quantitative activity. J. Immunol. 120:2027–2033.
37. Mochizuki, D. Y., Eisenman, J., Conlon, P. J., Park, D. L. 1986. Development and characterization of antis granulocyte-macrophage colony stimulating factor 136:3760–3709.
38. Prestidge, R. L., Watson, J. D., Urdal, D. L., Mochizu P., and Gillis, S. 1984. Biochemical comparison of stimulating factors secreted by a T cell lymphoma and cytic leukemia. J. Immunol. 133:295–298.
39. Grabstein, K., Eisenman, J., Mochizuki, D., Shanebeck, Hopp, T., March, C., and Gillis, S. 1986. Purification of B-cell stimulating factor: A molecule that stimulates multiple lymphokine-dependent cell lines. J. Exp. M 1414.
40. Metcalf, D. 1986. The molecular biology and functions macrophage colony stimulating factors. Blood 67:257–

About Entrez
Text Version
Entrez PubMed
Overview
Help | FAQ
Tutorial
New/Noteworthy
E-Utilities
PubMed Services
Journals Database
MeSH Database
Single Citation Matcher
Batch Citation Matcher
Clinical Queries
LinkOut
My NCBI (Cubby)
Related Resources
Order Documents
NLM Catalog
NLM Gateway
TOXNET
Consumer Health
Clinical Alerts
ClinicalTrials.gov
PubMed Central

Limits: **Publication Date to 1998**

Abstract

1: J Biol Chem. 1995 Sep 22;270(38):22507-13. Related Articles, Links

Characterization of the Phe-81 and Val-82 human fibroblast collagenase catalytic domain purified from Escherichia coli.

Gehring MR, Condon B, Margosiak SA, Kan CC.

Agouron Pharmaceuticals, Inc., San Diego, California 92121, USA.

Soluble recombinant human fibroblast collagenase catalytic domain was highly expressed and purified from Escherichia coli. The expression construct utilized the T7 gene 10 promoter for transcription of a two-cistron messenger RNA which encoded the ubiquitin-collagenase catalytic domain fusion protein as the second cistron. The ubiquitin domain was attached to the collagenase catalytic domain with the linker sequences Gly-Gly-Thr-Gly-Asp-Val-Ala-Gln (wild type) or Gly-Gly-Thr-Gly-Asp-Val-Gly-His (mutant) which served as cleavage sites for in vitro activation. The last four residues of the linker were included based on the crystal structure of human prostromelysin-1 catalytic domain. Soluble fusion proteins purified from E. coli retained the proteolytic activity of the collagenase catalytic domain. The collagenase catalytic domain was released by either autoproteolytic or stromelysin-1-catalyzed cleavage, purified to homogeneity, and separately possess Phe-81, Val-82, or Leu-83 as the amino-terminal residue. Very similar kcat/Km values were determined for the Phe-81 and Val-82 forms using continuous fluorogenic and chromogenic peptide cleavage assays.

PMID: 7673241 [PubMed - indexed for MEDLINE]

Limits: **Publication Date to 1998**

1: Gene. 1997 Jul 9;193(2):229-37. (abstract

High-level production and purification of biologically active proteins from bacterial and mammalian cells using the tandem pGFLEX expression system.

Manoharan HT, Gallo J, Gulick AM, Fahl WE.

McArdle Laboratory for Cancer Research, University of Wisconsin, Madison 53706, USA.

Because of the complexities involved in the regulation of gene expression in Escherichia coli and mammalian cells, it is considered general practice to use different vectors for heterologous expression of recombinant proteins in these host systems. However, we have developed and report a shuttle vector system, pGFLEX, that provides high-level expression of recombinant glutathione S-transferase (GST) fusion proteins in E. coli and mammalian cells. pGFLEX contains the cytomegaloma virus (CMV) immediate-early promoter in tandem with the E. coli lacZpo system. The sequences involved in gene expression have been appropriately modified to enable high-level production of fusion proteins in either cell type. The pGFLEX expression system allows production of target proteins fused to either the N or C terminus of the GST pi protein and provides rapid purification of target proteins as either GST fusions or native proteins after cleavage with thrombin. The utility of this vector in identifying and purifying a component of a multi-protein complex is demonstrated with cyclin A. The pGFLEX expression system provides a singular and widely applicable tool for laboratory or industrial production of biologically active recombinant proteins in E. coli and mammalian cells.

PMID: 9256081 [PubMed - indexed for MEDLINE]

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☑ BLACK BORDERS
- ☑ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☑ FADED TEXT OR DRAWING
- ☑ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☑ SKEWED/SLANTED IMAGES
- ☑ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☑ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☑ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.